

The Flow of Information in Interactive Quantum Protocols : the Cost of Forgetting *

Mathieu LAURIÈRE¹ and Dave TOUCHETTE^{2,3}

¹NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai,
email: mathieu.lauriere@gmail.com

²Institute for Quantum Computing, and Department of Combinatorics and Optimization,
University of Waterloo, email: touchette.dave@gmail.com

³Perimeter Institute for Theoretical Physics

Abstract

In the context of two-party interactive quantum communication protocols, we study a recently defined notion of quantum information cost (QIC), which possesses most of the important properties of its classical analogue, see Ref. [Tou15]. Notably, its link with amortized quantum communication complexity has been used in Ref. [BGKK⁺15] to prove an (almost) tight lower bound on the bounded round quantum complexity of Disjointness. However, the only known characterization of QIC was through a notion of purification of the input state. Although this definition has the advantage to be valid for fully quantum inputs and tasks, its interpretation for classical tasks remained rather obscure. Also, the link between this new notion and other notions of information cost for quantum protocols that had previously appeared in the literature (e.g. in Refs. [JRS03, JN14, KLLGR15]) was not clear, if existent at all.

We settle both these issues: for quantum communication with classical inputs, we provide an alternate characterization of QIC in terms of information about the input registers, avoiding any reference to the notion of a purification of the classical input state. We provide an exact operational interpretation of this alternative characterization as the sum of the cost of transmitting information about the classical inputs and the cost of forgetting information about these inputs. To obtain this characterization, we prove a general lemma, the Information Flow Lemma, assessing exactly the transfer of information in general interactive quantum processes. Specializing this lemma to interactive quantum protocols accomplishing classical tasks, we are also able to demystify the link between QIC and these other previous notions of information cost in quantum protocols. Furthermore, we clarify the link between QIC and IC of classical protocols by simulating quantumly classical protocols.

Finally, we apply these concepts to argue that any quantum protocol that does not forget information solves Disjointness on n -bits in $\Omega(n)$ communication, completely losing the quadratic quantum speedup. This provides a specific sense in which forgetting information is a necessary feature of interactive quantum protocols in order to obtain any significant improvement over classical protocols. We also apply these concepts to prove that QIC at zero-error is exactly n for the Inner Product function, and $n(1 - o(1))$ for a random Boolean function on $n + n$ bits.

*A one-page abstract of this work will appear in the Proceedings of the 8th Innovations in Theoretical Computer Science conference (ITCS 2017).

Contents

1	Introduction	1
2	Preliminaries: Quantum Communication and Information	3
3	Information Flow Lemma	6
4	Making Safe Copies of the Inputs	8
5	The Cost of Forgetting: a New Characterization of QIC	11
5.1	Alternate Definitions of Information Costs for Protocols with Classical Inputs	12
5.2	Operational Interpretation of HIC in Terms of CIC and CRIC	13
5.3	Operational Interpretation of QIC in Terms of CIC and CRIC	14
5.4	QIC and CIC are Almost Equivalent	15
5.5	Running Protocols on Superposition of Inputs	16
5.5.1	Product Distributions	16
5.5.2	General Distributions	17
6	Forgetting Information in Classical Protocols	19
6.1	Extending the Classical Setting : a New Characterization of IC	19
6.2	Reversible Classical Protocols	20
7	Disjointness: Speed-up for Quantum Protocols needs Forgetting Information	22
8	Quantum Simulation of Classical Protocols	25
9	Clean Protocols, IP, and Random Functions	28
9.1	Clean Protocols and Phase Encoding of the Output	28
9.2	Relating to $QIC(\Pi, \mu)$	29
9.3	Information Lower Bound	30
9.4	Inner Product function	32
9.5	Random Functions	32
9.6	Non-Zero Error and Classical Protocols	33
	References	33
A	Proofs for Section 7	35
A.1	Proof of Lemma 38	35
A.2	Proof of Lemma 39	35
B	The Various Notions of Information Cost	36

1 Introduction

Background. In two-party communication complexity [Yao79], Alice and Bob receive inputs x and y and run an interactive communication protocol by exchanging messages in order to compute $f(x, y)$ for some function f that depends on both these inputs. Their goal is to minimize the communication cost (denoted CC and QCC respectively in the classical and the quantum settings), that is, the amount of communication (bits or qubits). This model has found numerous applications in many areas of computer science. For excellent introductions to classical and quantum communication complexities, we refer the reader to [KN97] and [dW02] respectively.

One question that has received a lot of attention recently is whether it is possible to perform such protocols without leaking much information. In classical communication protocols, the information cost (IC) is defined as the information that the transcript reveals to each player about the input of the other one. In quantum communication protocols [Yao93], the registers are in a quantum state, which, in general, prevents the player from keeping track of the previous messages due to the no-cloning theorem. Nevertheless, the parties have quantum workspaces, where they may keep information about previous messages. The question is then to calculate how much information every new message reveals to them, given that they already know their own input and have kept some information in their quantum workspace according to the protocol.

Several notions of information cost for quantum protocols have already been used in the literature, see e.g. Refs [KNTSZ07, Kla02, JRS03, JRS09, JN14]. Each notion was somehow tailor-made for a specific purpose and very useful in that particular case. Nevertheless, these definitions did not seem to provide a general understanding of how information behaves in quantum communication. In Ref. [Tou15] has been introduced a general notion of Quantum Information Cost (QIC), which measures the total amount of *quantum* information about the inputs that is transmitted during the protocol. The corresponding notion of quantum information complexity of a function (the minimal QIC of a protocol computing the function) has been shown to exactly characterize the amortized communication complexity of that function, which is a fundamental property of the information complexity in the classical setting, see Ref. [BR11]. Moreover, this notion of QIC has already found multiple applications [Tou15, BGKK⁺15, NT16].

However, so far the only known characterization of QIC was through a notion of purification of the input state. Although this definition has the advantage to be valid for fully quantum inputs and tasks, its interpretation for classical tasks remained rather obscure. Also, the link between this new notion and other notions of information cost for quantum protocols that had previously appeared in the literature was not clear, if existent at all.

Our contributions. In this paper we shed a new light on the Quantum Information Cost (QIC), and settle both issues described above by relating this quantity to several other natural notions of information cost, including the classical IC, and by providing, when the inputs are classical, a new characterization of QIC which has an operational interpretation and does not require any reference to a purification register.

The cornerstone of our work is a general lemma, that we call the Information Flow Lemma (see Lemma 3), which precisely characterizes the transfer of information in quantum processes, run on arbitrary quantum inputs. This result then specializes to the setting we are interested in, namely quantum communication protocols. We stress that this lemma has already found other applications besides this work, in particular to prove a lower bound on quantum information complexity of the Augmented Index function on a uniform distribution over the zeros of the function [NT16], with corollaries on the space complexity of quantum streaming algorithm for the *DYCK*(2) problem of well-formed parentheses over two pairs of symbols.

We then turn our attention to quantum protocols with *classical inputs*. In this framework, even though some protocols might modify the input register, it is always possible, since the inputs are classical, to require that the players start the protocol by making a copy of their inputs and work with that copy. We call protocols such as these, where the input registers are left untouched, *safe protocols*. This seemingly

insignificant modification of the original protocol might drastically change the information cost. However, we prove that it can only decrease it (see Proposition 9). So it is enough to study the information cost of safe protocols when we are interested in minimizing the QIC for computing a task with classical inputs.

When studying such quantum protocols with classical inputs, a notion of information cost (called Classical input Information Cost, or CIC) has been introduced in Ref. [KLLGR15], where a first step was made to understand its relationship with QIC: the former is a lower bound on the latter – that is, $CIC \leq QIC$. In order to complete the picture, we introduce two new notions: the Holevo Information Cost (HIC), which measures how much information the players have about each other’s input *at the end* of the protocol (a round-by-round variant was considered in Ref. [JRS03, JN14]), and the Classical input Reverse Information Cost (CRIC), which counts how much information about the inputs is *forgotten* at each round by the player sending the message (this is somehow the dual under time reversal of CIC). Based on our Information Flow Lemma, we give new operational interpretations to these quantities and, informally speaking, we show that they satisfy the two following very natural relationships: the Holevo information cost corresponds to the amount of classical information that was learnt and not forgotten during the protocol, while the quantum information cost captures all of the information transmitted during the protocol (what was learnt plus what was forgotten). This yields a new characterization of QIC by CIC, up to a factor of 2. So the various notions of information cost introduced in this paper are tightly related, namely (see Propositions 16, 17 and 19):

Main Result 1: *We have: $HIC = CIC - CRIC$, $QIC = CIC + CRIC$. Moreover, $CIC \leq QIC \leq 2 \cdot CIC$.*

These relationships emphasize the importance of CRIC, the cost of forgetting information. This last quantity would always be zero in classical protocols: implicitly, classical information is always cloneable, hence players can memorize the whole history of the protocol and never forget information. To understand the link with quantum protocols forgetting information, we introduce a model of classical reversible computing, endowing classical protocols with the ability to forget information. We show that this feature can only increase their information cost, and, as such, forgetting information is somehow a wasteful phenomenon that should be avoided in the context of classical communication (see Theorem 33). However, in quantum protocols, cloning is not possible in general. This raises the question whether the property of forgetting information is only costly and should still be avoided in some sense. We answer this in the negative: forgetting information is absolutely necessary to obtain the quantum communication improvement allowed for computing certain functions. Indeed, if no information is forgotten in a quantum protocol, then $QIC = HIC$ is formally very similar to IC, and the continuity in the input distribution has no round dependence, as in the classical case. Thus, the round dependence in this continuity bound for general quantum protocols that do forget information [BGKK⁺15] can be understood as being due to the fact that the *same information* is forgotten and transmitted multiple times. With this observation, we prove that any quantum protocol for Disjointness that does not forget information has linear quantum communication complexity (see Theorem 36). Hence, quantum protocols that do not forget information cannot obtain the quadratic quantum speed-up for the Disjointness function [AA05], and this ability of quantum protocol to forget information is an essential feature of interactive quantum communication, not just some oddity we can get around. This can be summarized as follows:

Main Result 2 : *Forgetting information is useless in a classical reversible setting, but it is unavoidable in the quantum setting: it is a necessary feature of interactive quantum protocols to get significant communication improvement over classical protocols.*

This important distinction shows that the flow of information behaves quite differently in the classical and in the quantum setting. However, the classical communication complexity is always lower bounded by the quantum communication complexity: quantum messages can simulate classical ones. We can ask the same question in terms of information: is it always possible to quantumly simulate classical messages *while maintaining the information cost*? Our next main result provides a positive answer. We show that to any classical protocol Π_C corresponds a quantum simulation protocol Π_Q satisfying $QCC(\Pi_Q) = CC(\Pi_C)$, $QIC(\Pi_Q, \mu) = IC(\Pi_C, \mu)$ for any input distribution μ , and implementing the

same input-output channel $\Pi_Q = \Pi_C$. The main issue we deal with is the pure state quantum simulation of private randomness without altering the information cost (see Lemma 47).

Main Result 3 : *For any classical protocol, there exists a quantum protocol with the same input-output behaviour, and with communication and information costs smaller than the classical protocol.*

This result lets us conclude the paper with one more application. For the Inner Product function, QIC at zero-error over the uniform distribution is exactly n ; a similar lower bound of $n(1 - o(1))$ holds for a random Boolean function on $n + n$ bits. Further using the quantum simulation of classical protocols mentioned above together with the fact that classical IC is continuous at zero-error [BGPW13a], this shows that, in the limit when the error ε goes to 0, IC of such a random Boolean function is not only $\Omega(n)$ [BW12, KLL⁺15], but is precisely $n(1 - o(1))$ (such a tight bound for the IC of Inner Product was known from Ref. [BGPW13b]).

Outline of the paper. This paper is structured as follows. After some preliminaries (Section 2), we state and prove our Information Flow Lemma (Section 3). In Sections 4 and 5, we prove our results on safe quantum protocols, and then introduce CRIC, HIC and multiple other quantum notions of information cost (a table is provided in Appendix B to keep track of definitions and relationships). For the sake of comparison, in Section 6 we define IC in a classical reversible computation paradigm and show that forgetting information is wasteful. In contrast, we prove in Section 7 that there is no quantum communication speed-up for Disjointness when the quantum protocols are not allowed to forget information. Then, we show how to simulate quantumly classical protocols in Section 8. Finally we prove our results on Inner Product and random Boolean functions (Section 9).

2 Preliminaries: Quantum Communication and Information

Quantum Communication Model. Quantum communication complexity was introduced by Yao in Ref. [Yao93]. The model we use here is closer to the one of Cleve and Buhrman [CB97], with pre-shared entanglement, but we allow the players to communicate with quantum messages. In this model, an r -round protocol Π for a given classical task from input registers $A_{in} = X$, $B_{in} = Y$ to output registers A_{out} , B_{out} is defined by a sequence of isometries U_1, \dots, U_{M+1} along with a pure state $\psi \in \mathcal{D}(T_A^{in} T_B^{in})$ shared between Alice and Bob, for arbitrary finite dimensional registers T_A^{in} , T_B^{in} : the pre-shared entanglement. Above, $\mathcal{D}(\mathcal{A})$ is the set of all unit trace, positive semi-definite linear operators mapping \mathcal{A} into itself. See Refs [Wat15, Wil13]. We need $r + 1$ isometries in order to have r messages since a first isometry is applied before the first message is sent and a last one after the final message is received. In the case of even r , for appropriate finite dimensional quantum memory registers $A_1, A_3, \dots, A_{r-1}, A'$ held by Alice, $B_2, B_4, \dots, B_{r-2}, B'$ held by Bob, and quantum communication registers $C_1, C_2, C_3, \dots, C_r$ exchanged by Alice and Bob, we have $U_1 \in \mathcal{U}(A_{in} T_A^{in}, A_1 C_1)$, $U_2 \in \mathcal{U}(B_{in} T_B^{in} C_1, B_2 C_2)$, $U_3 \in \mathcal{U}(A_1 C_2, A_3 C_3)$, $U_4 \in \mathcal{U}(B_2 C_3, B_4 C_4)$, \dots , $U_r \in \mathcal{U}(B_{r-2} C_{r-1}, B_{out} B' C_r)$, $U_{r+1} \in \mathcal{U}(A_{r-1} C_r, A_{out} A')$, where $\mathcal{U}(\mathcal{A}, \mathcal{B})$ is the set of unitary channels from \mathcal{A} to \mathcal{B} : see Figure 1. We adopt the convention that, at the outset, $A_0 = A_{in} T_A^{in}$, $B_0 = B_{in} T_B^{in}$, for odd i with $1 \leq i < r$, $B_i = B_{i-1}$, for even i with $1 < i \leq r$, $A_i = A_{i-1}$ and also $B_r = B_{r+1} = B_{out} B'$, and $A_{r+1} = A_{out} A'$. In this way, after application of U_i , Alice holds register A_i , Bob holds register B_i and the communication register is C_i . In the case of an odd number of messages r , the registers corresponding to U_r, U_{r+1} are changed accordingly. We slightly abuse notation and also write Π to denote the channel from registers $A_{in} B_{in}$ to $A_{out} B_{out}$ implemented by the protocol, i.e. for any input distribution μ on XY and ρ_μ encoding μ on input registers $A_{in} B_{in}$,

$$\Pi(\rho_\mu) = \text{Tr}_{A' B'}(U_{M+1} U_M \dots U_2 U_1(\rho_\mu \otimes \psi)). \quad (2.1)$$

Note that the A' and B' registers are the final memory registers that are being discarded at the end of the protocol by Alice and Bob, respectively.

Recall that for a given state, all purifications are related by isometries on the purification registers. For classical input registers XY distributed according to μ , we consider a canonical purification $|\rho_\mu\rangle^{X R_X Y R_Y}$ of $\rho_\mu^{A_{in} B_{in}}$, with

$$|\rho_\mu\rangle^{X R_X Y R_Y} = \sum_{x,y} \sqrt{\mu(x,y)} |xxyy\rangle^{X R_X Y R_Y}. \quad (2.2)$$

We then say that the purifying registers $R_X R_Y$ contain *quantum copies* of XY . Then, the state at round i ,

$$\rho_i^{X R_X Y R_Y A_i B_i C_i} = U_i \dots U_1 (\rho^{X R_X Y R_Y} \otimes \psi^{T_A^{in} T_B^{in}}) \quad (2.3)$$

is pure. Also, we require that the final marginal state $\Pi(\rho^{A_{in} B_{in} R_X R_Y})$ on $R_X R_Y A_{out} B_{out}$ is classical. We say that a protocol Π solves a function f with error ε with respect to input distribution μ if $\Pr_\mu[\Pi(x,y) \neq f(x,y)] \leq \varepsilon$, and we say Π solves f with error ε if $\max_{(x,y)} \Pr[\Pi(x,y) \neq f(x,y)] \leq \varepsilon$.

We also make use of the notion of a control-isometry: it is an isometry acting on a classical-quantum register by leaving the content of the classical register unchanged. Such a classical register is called a control-register.

Quantum Information Cost. The main quantity of interest in this work is the quantum information cost, as introduced in [Tou15]. In quantum communication protocols, there is no clear notion of a transcript, so this definition counts how much information is exchanged in each round. In the sequel, we denote the Von Neumann entropy by H , and for a tripartite state ρ^{ABC} , we denote the conditional quantum mutual information (CQMI) between A and B conditioned on C by $I(A : B|C) = H(A, C) + H(B, C) - H(C) - H(A, B, C)$. We will make use of many properties of CQMI, among which the following.

Lemma 1 *If $\rho = \rho^{ABC}$ and $\sigma = \sigma^{DEF}$ are two states on distinct registers, then*

$$I(AD; BE|CF)_{\rho \otimes \sigma} = I(A; B|C)_\rho + I(D; E|F)_\sigma.$$

If $\rho = \rho^{ABCD} = \sum_c p(c) |c\rangle\langle c| \otimes \rho_c^{ABD}$ is a classical-quantum state with classical register C , then

$$I(A : B|CD)_\rho = \mathbb{E}_c [I(A : B|D)_{\rho_c}].$$

If $\rho = \rho^{ABCD}$ is a pure state, then

$$I(A; B|C)_\rho = I(A; B|D)_\rho.$$

Let us recall the definition of quantum information cost introduced in [Tou15].

Definition 2 *For a protocol Π and an input distribution μ , we define the quantum information cost of Π on input μ as*

$$\text{QIC}(\Pi, \rho) = \sum_{i \geq 1, \text{ odd}} I(C_i; R_X R_Y | B_i) + \sum_{i \geq 1, \text{ even}} I(C_i; R_X R_Y | A_i).$$

For any function f , any input distribution μ , and any $\varepsilon > 0$

$$\text{QIC}(f, \mu, \varepsilon) = \inf_{\Pi} \text{QIC}(\Pi, \mu) \quad (2.4)$$

where the infimum is over the protocols Π computing f with error ε w.r.t μ .

This quantity has many nice properties (see [Tou15, BGKK⁺15]); in particular it characterizes the (quantum) amortized communication complexity. We stress that the definition is independent of the choice of purification.

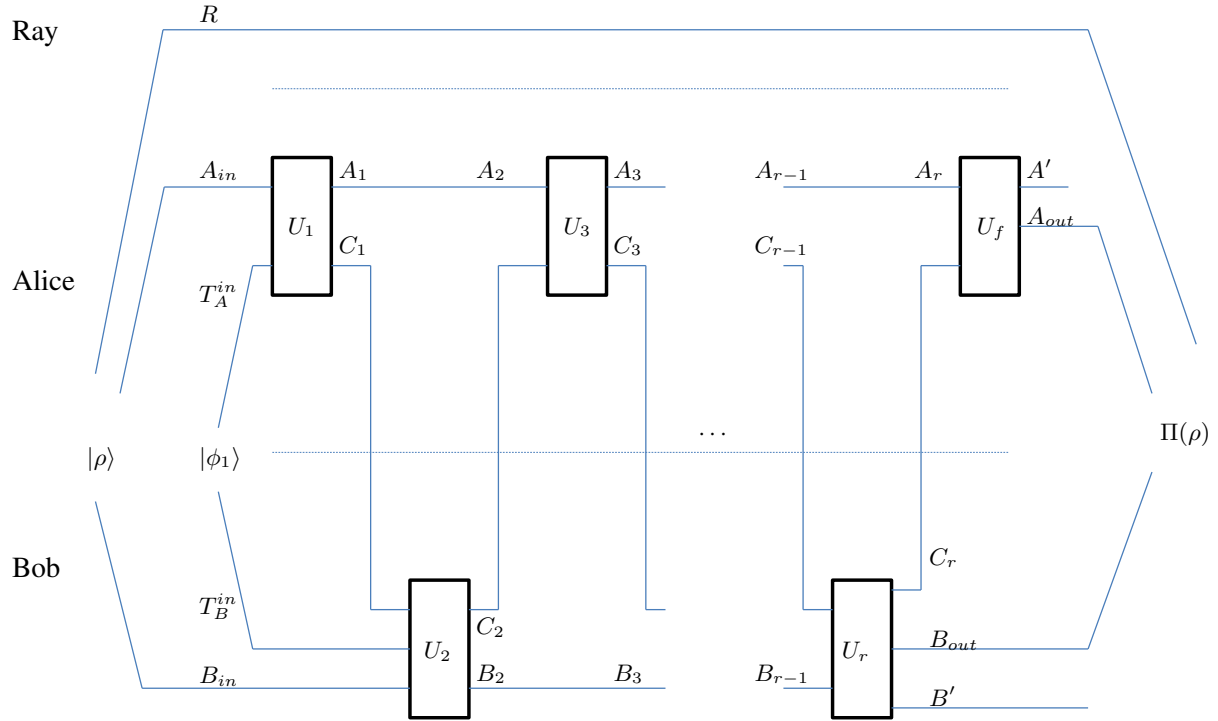


Figure 1: Depiction of a quantum protocol in the interactive model, adapted from the long version of [Tou15, Figure 1].

Discussion about compression. Some previous notions of information cost for quantum protocols (e.g. in Refs. [JRS03, JN14, KLLGR15]) were more similar in spirit to classical input information cost than to quantum information cost. Our results shed new light on why these previous definitions were restricted to compression results for a single round. In the first round, Alice does not yet possess any information on Bob’s input (aside from what she can infer from her own input). For one-round protocols, it is then immaterial whether one uses classical input information cost or quantum information cost. But then in subsequent rounds, generally Alice has in her registers some information about Bob’s input. It is then possible for her to forget information while sending a message. We can even construct a protocol where, at the third round, Bob does not learn anything whereas Alice forgets a lot of information. For such a round of communication, the previous definitions of information cost, e.g. CIC introduced in Ref. [KLLGR15], would evaluate to 0 whereas QIC would be large. Thus, it is impossible to compress such a quantum message down to its CIC, that is, almost at no cost, while keeping, in a round-by-round fashion, the overall state of the protocol almost equivalent to that in the original protocol. Indeed, we know from our developments that to forget information we must invest communication. As a consequence, we see that for quantum protocols, it is important to take into account the cost of forgetting information.

The purification register used in the definition of QIC possibly appears artificial when considering classical inputs. In this direction, we prove below (see Section 5) an arguably more natural characterization (at least from a classical correlation point of view) of each term in the quantum information cost as the sum of how much information about his own input a party is sending plus how much information about the other party’s input he is forgetting. However, we argue that there is still virtue in taking the purification of the classical input viewpoint. Firstly, it enables to keep track of a global pure state, which in many situations is a remarkably powerful viewpoint. Secondly and more fundamentally, the purification viewpoint has a nice operational interpretation through the task of quantum state redistribution, which is useful when aiming at compression results. Indeed, at any point of the interactive protocol, the pure quantum state can be seen as a 4-partite state $\rho^{A_R A_S M R}$ consisting of the receiver’s and the sender’s private registers (A_R and A_S respectively), the message register M and a purification register R . Then, each term in QIC is of the form $I(R; M|A_R)$, that is, the mutual information between the message and the inaccessible purification register, conditioned on the receiver’s side information. Such an expression is known [DY08, YD09] to quantify the cost of redistributing the message register while maintaining correlations with the receiver’s and the sender’s private registers as well as the environment. The CIC terms can also be given such an operational significance for the information about the sender’s input that a message contains. However, this viewpoint breaks down for the information that is forgotten (see the operational interpretation given at Section 5). Indeed, to measure the amount of information being forgotten, we condition on the sender’s side information for sending information about the receiver’s input. This term would be hard to account for in a compression viewpoint (unless we think of messages going backward). Hence, we think that the purification viewpoint remains appropriate for compression purposes.

3 Information Flow Lemma

In this section, we state and prove the *Information Flow Lemma* (see Lemma 3 below), which allows to keep track exactly of the flow of quantum information in an interactive protocol and is key to much of our further developments. Moreover, it gives a lower bound on QIC that does not depend on the number of round (see Corollary 5), and is used, among other things, to give an exact meaning to the cost of forgetting in interactive quantum protocols. We present here a quite general version of this result. However, we stress that a more limited version, that is still sufficient to obtain a lower bound on QIC, has already found some applications; see Ref. [NT16].

Let us consider the more general framework of bipartite interactive quantum processes, of which the model of quantum communication complexity defined in Section 2 is a special case. This general framework modelizes a discretized quantum process in which there is interaction between two distinct,

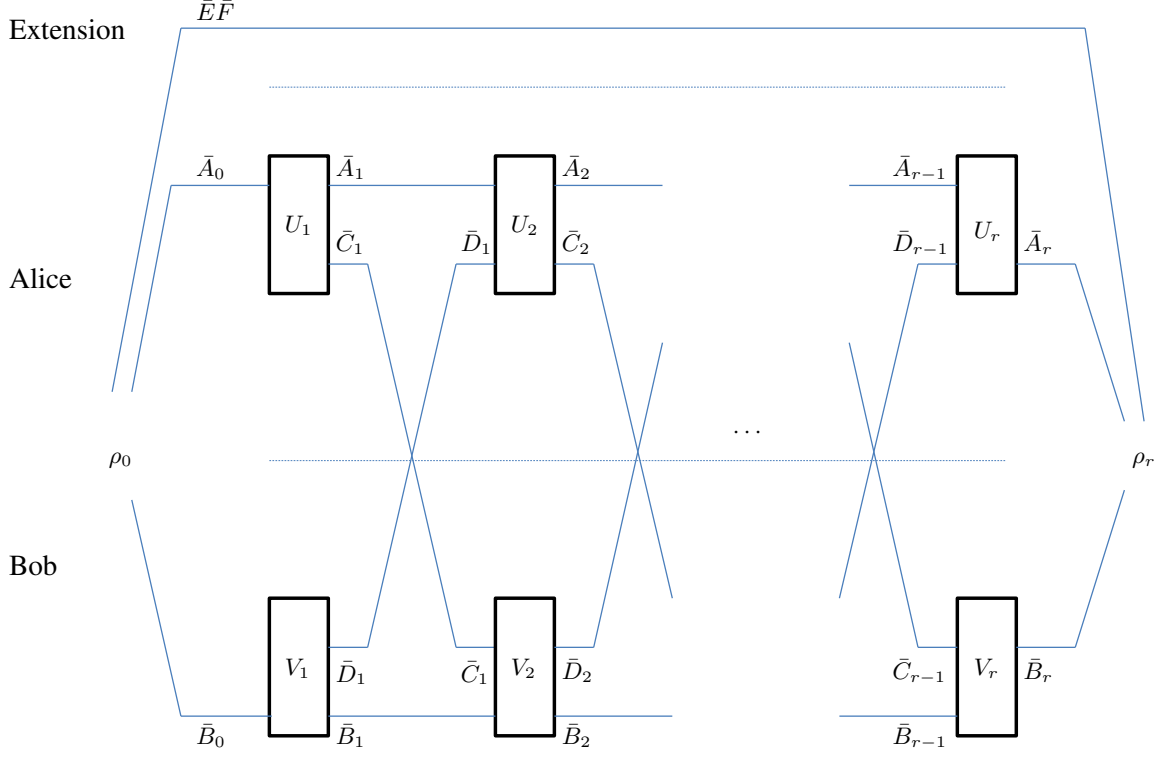


Figure 2: Depiction of an interactive quantum process, adapted from the long version of [Tou15, Figure 1].

localized parties, and local evolution at each time step.

In more details, Alice and Bob start in a joint state $\rho_0^{\bar{A}_0 \bar{B}_0}$, for which we consider an arbitrary extension $\rho_0^{\bar{A}_0 \bar{B}_0 \bar{E} \bar{F}}$ (such that $\text{Tr}_{\bar{E} \bar{F}}(\rho^{\bar{A}_0 \bar{B}_0 \bar{E} \bar{F}}) = \rho^{\bar{A}_0 \bar{B}_0}$). The process runs for $r + 1$ rounds, with ρ_i the state in round i , registers $\bar{A}_i, \bar{B}_i, \bar{C}_i$ and \bar{D}_i in each round, with $\bar{C}_0, \bar{D}_0, \bar{C}_{r+1}$ and \bar{D}_{r+1} being trivial registers in the 0-th and $r + 1$ -th round, initially and at the end of the process. In round i , for $1 \leq i \leq r$, after being generated by Alice, register \bar{C}_i gets communicated from Alice to Bob, and, after being generated by Bob, register \bar{D}_i gets communicated from Bob to Alice. Register \bar{A}_i is a quantum memory register held by Alice, and register \bar{B}_i is a quantum memory register held by Bob. The evolution is through local isometries $U_i = U_i^{\bar{A}_{i-1} \bar{D}_{i-1} \rightarrow \bar{A}_i \bar{C}_i}$ on Alice's side and $V_i = V_i^{\bar{B}_{i-1} \bar{C}_{i-1} \rightarrow \bar{B}_i \bar{D}_i}$ on Bob's side: $\rho_i^{\bar{A}_i \bar{B}_i \bar{C}_i \bar{D}_i \bar{E} \bar{F}} = (U_i \otimes V_i) \rho_{i-1}^{\bar{A}_{i-1} \bar{B}_{i-1} \bar{C}_{i-1} \bar{D}_{i-1} \bar{E} \bar{F}}$.

Registers $\bar{E} \bar{F}$ are left untouched throughout, and can be thought of in the following way: we want to measure how much information Bob knows about \bar{E} from the point of view of someone who knows \bar{F} . We get the following exact characterization of the flow of information from this point of view.

Lemma 3 (Information Flow Lemma) *Given an interactive quantum process as defined above, the fol-*

lowing holds:

$$I(\bar{E}; \bar{B}_{r+1}|\bar{F})_{\rho_{r+1}} - I(\bar{E}; \bar{B}_0|\bar{F})_{\rho_0} = \sum_{i=1}^r (I(\bar{E}; \bar{C}_i|\bar{F}\bar{B}_i)_{\rho_i} - I(\bar{E}; \bar{D}_i|\bar{F}\bar{B}_i)_{\rho_i}).$$

Proof.

We keep track of the flow of information using the chain rule and local isometric invariance of CQMI:

$$\begin{aligned} I(\bar{E}; \bar{B}_{r+1}|\bar{F}) &= I(\bar{E}; \bar{B}_r \bar{C}_r|\bar{F}) \\ &= I(\bar{E}; \bar{B}_r|\bar{F}) + I(\bar{E}; \bar{C}_r|\bar{F}\bar{B}_r) + (I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r) - I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r)) \\ &= (I(\bar{E}; \bar{B}_r|\bar{F}) + I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r)) + I(\bar{E}; \bar{C}_r|\bar{F}\bar{B}_r) - I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r) \\ &= I(\bar{E}; \bar{B}_r \bar{D}_r|\bar{F}) + I(\bar{E}; \bar{C}_r|\bar{F}\bar{B}_r) - I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r) \\ &= I(\bar{E}; \bar{B}_{r-1} \bar{C}_{r-1}|\bar{F}) + I(\bar{E}; \bar{C}_r|\bar{F}\bar{B}_r) - I(\bar{E}; \bar{D}_r|\bar{F}\bar{B}_r). \end{aligned}$$

Applying recursively the same argument leads to

$$\begin{aligned} I(\bar{E}; \bar{B}_{r+1}|\bar{F}) &= I(\bar{E}; \bar{B}_1 \bar{C}_1|\bar{F}) + \sum_{i=2}^r (I(\bar{E}; \bar{C}_i|\bar{F}\bar{B}_i) - I(\bar{E}; \bar{D}_i|\bar{F}\bar{B}_i)) \\ &= I(\bar{E}; \bar{B}_1 \bar{D}_1|\bar{F}) + I(\bar{E}; \bar{C}_1|\bar{F}\bar{B}_1) - I(\bar{E}; \bar{D}_1|\bar{F}\bar{B}_1) \\ &\quad + \sum_{i=2}^r (I(\bar{E}; \bar{C}_i|\bar{F}\bar{B}_i) - I(\bar{E}; \bar{D}_i|\bar{F}\bar{B}_i)) \\ &= I(\bar{E}; \bar{B}_0|\bar{F}) + \sum_{i=1}^r (I(\bar{E}; \bar{C}_i|\bar{F}\bar{B}_i) - I(\bar{E}; \bar{D}_i|\bar{F}\bar{B}_i)). \end{aligned}$$

We get the desired result by rearranging terms. ■

In the remainder of this work, we are concerned with quantum communication protocols as defined in Section 2, for which an easy corollary of the Information Flow Lemma is as follows. A similar result holds for Alice.

Corollary 4 *Given a protocol Π , an input distribution μ and any extension $\rho_0^{A_{in} B_{in} E_1 E_2}$ satisfying : $\text{Tr}_{E_1 E_2}(\rho_0^{A_{in} B_{in} E_1 E_2}) = \rho_\mu^{A_{in} B_{in}}$,*

$$I(E_1; B' B_{out}|E_2)_{\rho_{r+1}} - I(E_1; B_{in}|E_2)_{\rho_0} = \sum_{i \text{ odd}} I(E_1; C_i|E_2 B_i)_{\rho_i} - \sum_{i \text{ even}} I(E_1; C_i|E_2 B_i)_{\rho_i}.$$

Combining the above result and a similar one holding for Alice, we get the following lower bound on quantum information cost, stated as a sum of differences between the amount of correlations of reference registers with the output and the input.

Corollary 5 *Given a protocol Π , an input distribution μ and any two extensions $\rho_{0,B}^{A_{in} B_{in} E_1 E_2}, \rho_{0,A}^{A_{in} B_{in} F_1 F_2}$ satisfying : $\text{Tr}_{E_1 E_2}(\rho_{0,A}^{A_{in} B_{in} E_1 E_2}) = \rho_\mu^{A_{in} B_{in}}$, $\text{Tr}_{F_1 F_2}(\rho_{0,B}^{A_{in} B_{in} F_1 F_2}) = \rho_\mu^{A_{in} B_{in}}$, the following holds:*

$$\begin{aligned} \text{QIC}(\Pi, \rho) &\geq I(F_1; A_{out} A'|F_2) - I(F_1; A_{in}|F_2) \\ &\quad + I(E_1; B_{out} B'|E_2) - I(E_1; B_{in}|E_2). \end{aligned}$$

4 Making Safe Copies of the Inputs

In this section, we show that making safe copies of classical inputs at the outset of a quantum protocol never increases its quantum information cost. So, when studying the quantum information complexity of a function, it is always possible to assume that protocols do not change the input registers.

Following Ref. [JRS03], we introduce the notion of safe copies and safe protocols.

Definition 6 (Safe protocol) Recall that, in a quantum communication protocol implementing a classical task, players receive initial classical data in some quantum input registers. We say that such a protocol is safe if the players only use these input registers as control registers.

Note that for quantum protocols, making a local copy of the classical input does not change the quantum communication cost. However it is not obvious from definition that the same property should be true for the information cost. Let us make this question more precise by associating to every protocol another protocol, which is safe.

Safe Version of a Protocol. Consider any protocol Π . We define a safe version of Π as follows. Let Π' be the protocol in which Alice and Bob first make a coherent (safe) copy of their respective inputs X, Y at the outset of the protocol into safe registers X', Y' , and then run Π while using X' and Y' as inputs. Recall that there are also coherent copies held in purification registers R_X, R_Y . That is, on input distribution μ , we denote as ρ_μ^{XY} the state

$$\rho_\mu^{XY} = \sum_{x,y} \mu(x,y) |x\rangle\langle x|^X \otimes |y\rangle\langle y|^Y, \quad (4.1)$$

and we consider a purification of the form

$$|\rho_\mu\rangle^{XYR_XR_Y} = \sum_{x,y} \sqrt{\mu(x,y)} |x\rangle^X |y\rangle^Y |x\rangle^{R_X} |y\rangle^{R_Y}. \quad (4.2)$$

In the protocol Π' , the registers X, Y are then left untouched for the remainder of the protocol, which is identical to protocol Π acting on input registers X', Y' after such copies are made. We want to show that the quantum information cost of Π' is never greater than that of Π . More formally, define the isometries

$$U_X^{X \rightarrow XX'} = \sum_{x \in X} |x\rangle^X |x\rangle^{X'} \langle x|^X, \quad (4.3)$$

$$U_Y^{Y \rightarrow YY'} = \sum_{y \in Y} |y\rangle^Y |y\rangle^{Y'} \langle y|^Y. \quad (4.4)$$

Then the safe protocol Π' is defined from Π by:

1. applying U_X and then U_1 acting on X' on Alice's side in the first round,
2. applying U_Y and then U_2 acting on Y' on Bob's side in the second round,
3. running U_i in round i for $i \geq 3$.

This does not change the classical input/output behavior of the protocol. If we think of acting U_Y before U_1 , this does not change the value of any QIC term, and we get state

$$|\rho'_\mu\rangle^{XX'R_XYY'R_Y} = (U_X^{X \rightarrow XX'} \otimes U_Y^{Y \rightarrow YY'}) |\rho_\mu\rangle^{XR_XYR_Y} \quad (4.5)$$

$$= \sum_{x,y} \sqrt{\mu(x,y)} |xxxyyy\rangle^{XX'R_XYY'R_Y} \quad (4.6)$$

at the outset of protocol Π' . We then show that making such safe copies does not increase the QIC of a protocol.

Making Safe Copies can only Decrease QIC of a Protocol. It turns out that $\text{QIC}(\Pi, \mu)$ and $\text{QIC}(\Pi', \mu)$ can be very different. Let us illustrate this point with a simple example.

Example 7 Consider an input distribution μ such that X is uniformly distributed, and $Y = X$. Consider a protocol in which Alice directly sends her input to Bob. Then the costs are

$$\text{QIC}(\Pi, \mu) = I(X : R_X R_Y | Y)_{\rho_\mu} \quad (4.7)$$

$$= I(X : R_X R_Y)_{\rho_\mu} \quad (4.8)$$

$$= H(X)_{\rho_\mu} \quad (4.9)$$

$$= \lg |X|, \quad (4.10)$$

$$\text{whereas } \text{QIC}(\Pi', \mu) = I(X' : R_X R_Y | Y' Y)_{\rho'_\mu} \quad (4.11)$$

$$= 0, \quad (4.12)$$

in which we used for $\text{QIC}(\Pi', \mu)$ that all registers are classical once X is traced out along with the fact that $X = Y$, similarly for $I(X : R_X R_Y)_{\rho_\mu}$ and tracing out Y , and finally, since $\rho_\mu^{X Y R_X R_Y}$ is pure, $I(X : R_X R_Y | Y)_{\rho_\mu} = I(X : R_X R_Y)_{\rho_\mu}$.

This phenomenon might occur even when there is no correlation between X and Y , as shown by the following example.

Example 8 Consider an input distribution μ such that X and Y are distributed independently and uniformly. Consider a protocol in which Alice directly sends her input to Bob. Then the costs are

$$\text{QIC}(\Pi, \mu) = I(X : R_X R_Y | Y)_{\rho_\mu} \quad (4.13)$$

$$= I(X; R_X)_{\rho_\mu} \quad (4.14)$$

$$= 2H(X)_{\rho_\mu} \quad (4.15)$$

$$= 2 \lg |X|, \quad (4.16)$$

$$\text{whereas } \text{QIC}(\Pi', \mu) = I(X' : R_X R_Y | Y' Y)_{\rho'_\mu} \quad (4.17)$$

$$= I(X'; R_X)_{\rho'_\mu} \quad (4.18)$$

$$= H(X')_{\rho'_\mu} \quad (4.19)$$

$$= \lg |X|, \quad (4.20)$$

where we used that $\rho_\mu^{X R_X}$ is a pure state whereas ρ'_μ is classical on $X' R_X$ once X is traced out.

One can check that, if Bob sends register X back to Alice (without copying it), $\text{QIC}(\Pi, \mu)$ increases to $4 \lg |X|$ while $\text{QIC}(\Pi', \mu)$ increases to $2 \lg |X|$ only. Moreover, if Bob first makes a copy of X before sending it back, $\text{QIC}(\Pi, \mu)$ increases to $3 \lg |X|$ while $\text{QIC}(\Pi', \mu)$ stays at $\lg |X|$. By repeating this process for r rounds, $\text{QIC}(\Pi, \mu)$ increases to $(2r + 1) \lg |X|$ while $\text{QIC}(\Pi', \mu)$ stays at $\lg |X|$, and we can make these information costs as different as we like.

The examples above show that making safe copies might influence a lot the quantum information cost. However, we show that this operation can only decrease QIC.

Proposition 9 For any protocol Π and any input distribution μ for X, Y , the safe version of Π , the protocol Π' defined above, satisfies

$$\text{QIC}(\Pi', \mu) \leq \text{QIC}(\Pi, \mu). \quad (4.21)$$

Moreover, if Π is already a safe protocol, then we have equality.

Proof. Before running protocol Π , let us first relabel the classical inputs X, Y as X', Y' , and then apply $U_X^{R_X \rightarrow R_X X}$ and $U_Y^{R_Y \rightarrow R_Y Y}$ on R_X, R_Y in order to recreate coherent copies of the input in registers X, Y . The state at this point is then the same as in Π' before starting to apply the U_i 's (if we think of applying U_Y on Bob's side before U_1 on Alice's side, which does not change the information cost), since that protocol is invariant under how the additional coherent copy of X and Y is created. If we then

run Π using the coherent copies in registers X', Y' as inputs, the state in each round is then the same as in Π' . Notice that up to relabeling of the input registers and application of the isometries on R_X, R_Y , the protocol just defined is equivalent to Π , and hence it has the same information cost, with terms $I(R_X R_Y; C_i | B_i)_{\rho_i} = I(R_X R_Y X Y; C_i | B_i)_{\rho'_i}$ in round i , in contrast to the information cost terms in Π' , which are of the form $I(R_X R_Y; C_i | Y B_i)_{\rho'_i}$. The result follows since for each i ,

$$I(R_X R_Y X Y; C_i | B_i)_{\rho'_i} = I(Y; C_i | B_i)_{\rho'_i} + I(R_X R_Y; C_i | Y B_i)_{\rho'_i} + I(X; C_i | R_X R_Y Y B_i)_{\rho'_i} \quad (4.22)$$

$$\geq I(R_X R_Y; C_i | Y B_i)_{\rho'_i}, \quad (4.23)$$

and the terms $I(Y; C_i | B_i)_{\rho'_i}$ and $I(X; C_i | R_X R_Y Y B_i)_{\rho'_i} = I(X; C_i | A_i)_{\rho'_i}$ vanish whenever Π is a safe protocol, holding throughout an unmodified copy of X' in A_i and of Y' in B_i . The result follows. ■

As a consequence, whenever we are interested in minimizing the quantum information cost, we may always consider such protocols that start by making a local copy of their inputs. This implies the following for the quantum information complexity of a function :

Corollary 10 *For any function f , any input distribution μ , and any $\varepsilon > 0$*

$$\text{QIC}(f, \mu, \varepsilon) = \inf_{\Pi'} \text{QIC}(\Pi', \mu), \quad (4.24)$$

where the infimum is over the safe protocols Π' computing f with error ε w.r.t μ .

Note that here, in contrast with (2.4), the minimum is over a smaller class of protocols. In the sequel, unless otherwise specified, we only consider safe protocols.

5 The Cost of Forgetting: a New Characterization of QIC

In this section, we show that even though quantum protocols are reversible and thus can somehow forget information, there is a quantum information cost associated in particular with forgetting classical information. The fact, proven in the previous section, that unsafe protocols might have higher information cost than their safe counterpart can be seen as an example of this phenomenon for a party forgetting information about his own input. We focus here on safe protocols and consider the cost of forgetting information learnt previously about the other party's input. The remark at the end of Example 8 can be thought of as a simple, avoidable occurrence of this phenomenon. We will see later that in general for quantum protocols, it is not always possible to avoid this cost of forgetting information.

We introduce the Holevo Information Cost, defined as the amount of information the players have at the end of the protocol. We show that it is exactly characterized as the amount of information learnt minus the amount of information forgotten. This relation even holds at any intermediate stage of the protocol. We also consider how much Holevo information a party can obtain if he runs (part of) his input in superposition.

Note that the information flow lemma, characterizing exactly the flow of quantum information in interactive protocols, can be seen as a fully quantum generalization of this result.

For protocols with classical inputs, we provide an alternative characterization of their quantum information cost that does not require introducing a purification register. More precisely, we show that at each round, QIC can be divided into two parts: the first one measures how much information is sent by one party to the other one; the second one counts how much information the party sending the message is forgetting about the other party's input. This additional term does not exist in classical communication because players can always keep copies of all past messages, so they never forget information. But in quantum communication, cloning is in general impossible and players cannot always keep all the information they have received.

5.1 Alternate Definitions of Information Costs for Protocols with Classical Inputs

We first recall the notion of classical input information cost introduced by Kerenidis, Laurière, Le Gall and Rennela in [KLLGR15, KLLGR16]. They also define an asymmetric version of quantum information cost. They have the following definitions, in which we consider safe protocols and split Alice's local register in round i as XA_i and similarly as YB_i for Bob.

Definition 11 *For a protocol Π and an input distribution μ , the classical input information cost of the messages from Alice to Bob (resp. from Bob to Alice) is defined as*

$$\begin{aligned} \text{CIC}_{A \rightarrow B}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; X|YB_i) \\ (\text{resp. } \text{CIC}_{B \rightarrow A}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ even}} I(C_i; Y|XA_i)), \end{aligned}$$

and the quantum information cost of the messages from Alice to Bob (resp. from Bob to Alice) as

$$\begin{aligned} \text{QIC}_{A \rightarrow B}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; R_X R_Y | YB_i) \\ (\text{resp. } \text{QIC}_{B \rightarrow A}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ even}} I(C_i; R_X R_Y | XA_i)). \end{aligned}$$

It follows from the data processing inequality that CIC is always at most QIC.

Proposition 12 ([KLLGR15, KLLGR16]) *For any protocol Π and any input distribution μ ,*

$$\text{CIC}_{A \rightarrow B}(\Pi, \mu) \leq \text{QIC}_{A \rightarrow B}(\Pi, \mu), \quad \text{CIC}_{B \rightarrow A}(\Pi, \mu) \leq \text{QIC}_{B \rightarrow A}(\Pi, \mu). \quad (5.1)$$

Note that $\text{QIC}(\Pi, \mu) = \text{QIC}_{A \rightarrow B}(\Pi, \mu) + \text{QIC}_{B \rightarrow A}(\Pi, \mu)$, so we define similarly a symmetric version of classical input information cost of the protocol Π as

$$\text{CIC}(\Pi, \mu) = \text{CIC}_{A \rightarrow B}(\Pi, \mu) + \text{CIC}_{B \rightarrow A}(\Pi, \mu). \quad (5.2)$$

We want to compare these two quantities, and in particular we find that they are related with a further notion of information cost, which we call the Holevo information cost. This quantity evaluates the Holevo information each party possesses at the end of the protocol about the other party's input, conditional on his own input.

Definition 13 *For a protocol Π and an input distribution μ , the Holevo information cost from Alice to Bob is defined as*

$$\text{HIC}_{A \rightarrow B}(\Pi, \mu) = I(X; B_{\text{out}} B' | Y),$$

and the Holevo information cost from Bob to Alice as

$$\text{HIC}_{B \rightarrow A}(\Pi, \mu) = I(Y; A_{\text{out}} A' | X).$$

We also define the (total) Holevo information cost as $\text{HIC}(\Pi, \mu) = \text{HIC}_{A \rightarrow B}(\Pi, \mu) + \text{HIC}_{B \rightarrow A}(\Pi, \mu)$.

Note that similar considerations can be made in each round i by considering the protocol Π_i that runs Π up to round i and then stops (with an appropriate partition of the registers in round i , depending on whether i is even or odd, and who holds C_i). For instance, in any odd round i , after reception by Bob of message C_i from Alice, the conditional Holevo information Bob has about Alice's input is: $I(X : B_i C_i | Y)$. Such variants appeared, e.g., in Refs [JRS03, JN14].

5.2 Operational Interpretation of HIC in Terms of CIC and CRIC

The quantity HIC corresponds to the information remaining at the end of the protocol. However, since in a quantum protocol it might be unavoidable to forget information along the way (because cloning is in general impossible), we cannot just count the information that was received: we should also quantify the amount of information that each player forgets. We introduce the following notion to take this phenomenon into account.

Definition 14 *For a protocol Π and an input distribution μ , the classical input reverse information cost of the messages from Bob back to Alice (resp. from Alice to Bob) is defined as*

$$\begin{aligned} \text{CRIC}_{A \leftarrow B}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ even}} I(C_i; X | Y B_i) \\ \left(\text{resp. } \text{CRIC}_{B \leftarrow A}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; Y | X A_i) \right). \end{aligned}$$

We also define the total classical input reverse information cost of protocol Π as

$$\text{CRIC}(\Pi, \mu) = \text{CRIC}_{A \leftarrow B}(\Pi, \mu) + \text{CRIC}_{B \leftarrow A}(\Pi, \mu).$$

We soon make the above intuition more precise by providing an operational interpretation, but let us first consider a simple example.

Example 15 *Let μ be an input distribution with X, Y distributed independently and uniformly on n bits, and consider a protocol in which, after the second round, Alice has received a copy of Bob's input, Y . At this point, Alice copies the first m out of the n bits of Y , and sends back Y to Bob. Then the term with $i = 3$ in $\text{CRIC}_{B \leftarrow A}$ will amount to the $(n - m)$ bits of information about Y that Alice is forgetting.*

We now suggest an operational interpretation of CIC and CRIC. We can consider the following scenario. Let us fix a protocol Π . Consider a classical input state on registers XY purified in registers $R_X R_Y$. Alice is given her input X as usual, but also the purification R_Y of Bob's input. Bob is only given his input Y , and so only the register R_X is held in some reference register inaccessible to the both parties. Alice is given the register R_Y in order for her to be able to generate any state on $A_i B_i C_i$ in the protocol, for i odd as well as i even, and then transmit the message on C_i to Bob, after giving him his side information B_i . We are interested in how much new information about X this message C_i contains, hence we are only putting R_X in the referee's hand. More formally, suppose that we are interested in this information for round i . We then ask what is the asymptotic quantum communication cost for redistributing the C_i register of this state from Alice to Bob if, apart from C_i , Alice holds the A_i, X, R_Y registers and Bob holds the B_i, Y registers. This is $I(C_i : R_X | B_i Y) = I(C_i : X | Y B_i)$, for classical registers X, Y . Depending on whether i is odd or even, this is the i th term in $\text{CIC}_{A \rightarrow B}$ or in $\text{CRIC}_{A \leftarrow B}$ of the protocol Π (in the usual scenario where Alice does not have access to R_Y). Remember that quantum communication in state redistribution is symmetric under time-reversal [DY08, YD09], so that the cost is the same if Bob decides to send back this message to Alice. Hence, not only does this scenario gives an operational interpretation to CIC as the amount of information about X Alice is sending to Bob in odd rounds, but also to CRIC as the amount of information about X Bob is forgetting by sending it back to Alice in even rounds.

This interpretation leads to the following formal result.

Proposition 16 *Let μ be a distribution and Π be a safe protocol with classical inputs distributed according to μ . Then*

$$\begin{aligned} \text{HIC}_{A \rightarrow B}(\Pi, \mu) &= \text{CIC}_{A \rightarrow B}(\Pi, \mu) - \text{CRIC}_{A \leftarrow B}(\Pi, \mu), \\ \text{HIC}_{B \rightarrow A}(\Pi, \mu) &= \text{CIC}_{B \rightarrow A}(\Pi, \mu) - \text{CRIC}_{B \leftarrow A}(\Pi, \mu), \\ \text{HIC}(\Pi, \mu) &= \text{CIC}(\Pi, \mu) - \text{CRIC}(\Pi, \mu). \end{aligned}$$

Proof of Proposition 16.

From the above operational interpretation of CIC and CRIC, it is then intuitive that in any odd round i , after reception by Bob of message C_i from Alice, the conditional Holevo information $I(X' : B_i C_i | Y')$ Bob has about Alice's input can be written as follows:

$$I(X : B_i C_i | Y) = \sum_{j \text{ odd } j \leq i} I(C_j : X | Y B_j) - \sum_{j \text{ even } j \leq i} I(C_j : X | Y B_j), \quad (5.3)$$

in which on the right hand side the first sum corresponds to terms in $\text{CIC}_{A \rightarrow B}$ and the second one to terms in $\text{CRIC}_{A \leftarrow B}$. Note that this equality follows from Corollary 4, direct consequence of the Information Flow Lemma, with classical extension registers $E_1 = X$, $E_2 = Y$ (classical copies of these registers), along with the fact that for two classical copies Y_1, Y_2 of Y , $I(C_i; X | Y_1 Y_2 B_i) = I(C_i; X | Y B_i)$, $I(X; Y_1 B_i C_i | Y_2) = I(X; B_i C_i | Y_2)$, and $I(X; Y_1 | Y_2) = 0$. If r is odd, $I(X; B_r C_r | Y) = I(X; B_{\text{out}} B' | Y)$ and the result follows. If r is even, $I(X; B_r C_r | Y) = I(X; B_{\text{out}} B' C_r | Y) = I(X; B_{\text{out}} B' | Y) + I(X; C_r | Y B_r)$. Similar statements hold for Alice, with the role of odd and even rounds interchanged. The statement follows. ■

5.3 Operational Interpretation of QIC in Terms of CIC and CRIC

The introduction of the reference register R in the definition of quantum information cost, which can be decomposed into $R = R_X, R_Y$ for classical inputs, is natural when discussing compression while keeping quantum correlations, and for general quantum inputs. But when discussing protocols implementing classical tasks it might appear somewhat artificial. We now present an alternative characterization of quantum information cost on classical inputs that does not involve such purification registers and only mention the classical input registers, similar to the notion of classical input information cost (CIC) of Ref. [KLLGR15, KLLGR16]. We start by expanding the i th term in the quantum information cost. For odd i ,

$$I(C_i : R_X R_Y | Y B_i)_{\rho'_i} = I(C_i : R_X | Y B_i)_{\rho'_i} + I(C_i : R_Y | R_X Y B_i)_{\rho'_i} \quad (5.4)$$

(we could do similarly for even i with the conditioning instead on $X' A_i$). The first term on the right hand side is the classical input information cost term $I(C_i; R_X | Y B_i) = I(C_i; X | Y B_i)$ in round i and somehow quantifies the amount of information that message C_i contains about X for someone who already knows Y and possesses B_i as quantum side-information, while the second one does not immediately have such an intuitive interpretation. However, we can rewrite it as $I(C_i : R_Y | X A_i) = I(C_i : Y | X A_i)$ since $X A_i$ contain a purification of $\rho_i^{B_i C_i R_X R_Y Y}$. Notice that X, Y are both classical in this term, which can now be informally interpreted as the amount of information that message C_i contains about Y for someone who already knows X and possess A_i . But remember that it is Alice who generated message C_i , so in a classical protocol A_i would contain a copy of C_i and this term would always evaluate to 0. However, quantum protocols are reversible, so it is somehow possible to forget information along the way. This term then corresponds, in a sense made precise by Proposition 16, to the amount of information Alice is forgetting about Y when transmitting C_i (CRIC).

This leads to the following result.

Proposition 17 *Let μ be a distribution and Π be a safe protocol with classical inputs distributed according to μ . Then*

$$\begin{aligned} \text{QIC}_{A \rightarrow B}(\Pi, \mu) &= \text{CIC}_{A \rightarrow B}(\Pi, \mu) + \text{CRIC}_{B \leftarrow A}(\Pi, \mu), \\ \text{QIC}_{B \rightarrow A}(\Pi, \mu) &= \text{CIC}_{B \rightarrow A}(\Pi, \mu) + \text{CRIC}_{A \leftarrow B}(\Pi, \mu), \\ \text{and} \quad \text{QIC}(\Pi, \mu) &= \text{CIC}(\Pi, \mu) + \text{CRIC}(\Pi, \mu). \end{aligned}$$

5.4 QIC and CIC are Almost Equivalent

We show that, even though the asymmetric versions of QIC and CIC can be very different as exhibited in Ref. [KLLGR15, KLLGR16], the symmetric versions can only be separated by at most a factor of two. This can be understood intuitively by the fact that a protocol cannot forget more information than it transmits.

Theorem 18 *For any protocol Π and any input distribution μ , it holds that*

$$\text{CIC}(\Pi, \mu) \leq \text{QIC}(\Pi, \mu) \leq 2 \cdot \text{CIC}(\Pi, \mu).$$

Hence for any function f , any input distribution μ and any error threshold ε ,

$$\text{CIC}(f, \mu, \varepsilon) \leq \text{QIC}(f, \mu, \varepsilon) \leq 2 \cdot \text{CIC}(f, \mu, \varepsilon).$$

It was already noticed in Ref. [KLLGR15, KLLGR16], that $\text{CIC}(\Pi, \mu) \leq \text{QIC}(\Pi, \mu)$. So to prove the above result, it is sufficient to show the following.

Proposition 19 *For any protocol Π and any input distribution μ , it holds that*

$$\text{QIC}(\Pi, \mu) \leq 2 \cdot \text{CIC}(\Pi, \mu).$$

The proof relies on the characterization of the Holevo information cost given by Proposition 16.

Proof. We have:

$$\begin{aligned} \text{QIC}(\Pi, \mu) &= \text{CIC}(\Pi, \mu) + \text{CRIC}(\Pi, \mu) \\ &\leq \text{CIC}(\Pi, \mu) + \text{CRIC}(\Pi, \mu) + \text{HIC}(\Pi, \mu) \\ &= 2\text{CIC}(\Pi, \mu), \end{aligned} \tag{5.5}$$

where the inequality comes from the nonnegativity of Holevo information cost, that is $\text{HIC}(\Pi, \mu) \geq 0$, and the last equality holds by Proposition 16. ■

Since we believe that Proposition 19 helps understanding QIC better and might lead to new results involving this quantity, we provide an alternative proof sketch with a slightly different point of view. In particular, the symmetry of QIC with respect to a message being transmitted forward or backward is made evident, whereas the link between CIC and CRIC under such a reversal of direction for message transmission is also highlighted.

Alternative Proof Sketch of 15QIC and CIC are Almost Equivalenttheorem.19 19. Given a r -message protocol Π , let Π' be the protocol that runs Π forward but does not discard A' , B' , and then, without making any copy of the output, runs Π backward. Then, for any $k \in \{0, \dots, r-1\}$, the $(r+k)$ th message in Π' is identical to the $(r-k+1)$ th message, except that the roles of the sender and receiver have been exchanged. Since the terms in QIC are symmetric under time-reversal, we have $\text{QIC}_{A \rightarrow B}(\Pi', \mu) = \text{QIC}_{B \rightarrow A}(\Pi', \mu) = \text{QIC}(\Pi, \mu)$. So the CIC for Alice and Bob in Π' is respectively

$$\begin{aligned} \text{CIC}_{A \rightarrow B}(\Pi', \mu) &= \text{CIC}_{A \rightarrow B}(\Pi, \mu) + \text{CRIC}_{A \leftarrow B}(\Pi, \mu) = \text{QIC}(\Pi, \mu) \\ \text{and } \text{CIC}_{B \rightarrow A}(\Pi', \mu) &= \text{CIC}_{B \rightarrow A}(\Pi, \mu) + \text{CRIC}_{B \leftarrow A}(\Pi, \mu) = \text{QIC}(\Pi, \mu), \end{aligned}$$

since the last M messages in Π' consist of the M messages of Π run backward and thus the CIC of these messages in Π' correspond to the CRIC of Π . Thus, $\text{QIC}(\Pi', \mu) = 2 \cdot \text{QIC}(\Pi, \mu)$ and $\text{CIC}(\Pi', \mu) = \text{QIC}(\Pi, \mu)$. By (5.3) and the nonnegativity of Holevo information, $\text{CRIC}_{A \leftarrow B}(\Pi, \mu)$ is at most $\text{CIC}_{A \rightarrow B}(\Pi, \mu)$ and $\text{CRIC}_{B \leftarrow A}(\Pi, \mu)$ is at most $\text{CIC}_{B \rightarrow A}(\Pi, \mu)$, since it should not be possible to send back more information about the other party's input than what was received. This intuition also leads to the inequality $\text{QIC}(\Pi, \mu) \leq 2 \cdot \text{CIC}(\Pi, \mu)$. ■

In 15QIC and CIC are Almost Equivalenttheorem.19 19 we prove that QIC and CIC can be different by at most a factor of 2. In fact, one can see that a necessary and sufficient condition to have $\text{QIC} =$

CIC is that $\text{CRIC}(\Pi, \mu) = 0$, and then also $\text{QIC} = \text{HIC}$. Intuitively, this means that at each round the player who sends the message does not forget anything about what she has learnt in the previous rounds. Protocols with only a single message satisfy this property. Also, quantum simulation of classical protocols also satisfy this property; see Section 8.

At the other extreme, one can see that a sufficient condition to have $\text{QIC} = 2 \cdot \text{CIC}$ is that $\text{HIC}(\Pi, \mu) = 0$, which only happens if the protocols completely uncompute any information about its input (apart possibly locally encoded information, or, as we will discuss later, “phase” or “superposition” information). Nevertheless, this bound should be almost achieved by memoryless protocols (*i.e.* protocols using only input registers together with a pure message register C_i , and no private working space registers A_i, B_i). Say the message register C_r ends up with Bob, then $\text{QIC}(\Pi, \mu) = \text{CIC}(\Pi, \mu) - I(X; C_r|Y)$. However, players could also forget information much later than they learn it, and so memoryless protocols are not the only type of protocols achieving this bound.

5.5 Running Protocols on Superposition of Inputs

In the previous section, we considered the amount of information a party learnt and forgot about the other party’s classical input, when considering that he was also running on a classical input. However, in certain contexts, such as settings with privacy concerns [CVDNT99, K1a02, JRS09, KLLGR15, SSS15], other variants of the amount of information learnt by a party about the other party’s classical input are natural to consider, like the one corresponding to allowing that party to run on a quantum superposition of its intended input distribution. This makes for a quantum variant of the honest-but-curious classical paradigm, in which the party generates the correct “distribution over messages”, but wishes to learn as much information as possible while doing so.

5.5.1 Product Distributions

With this in mind, we now define an alternative notion of quantum information cost for product distributions, and a corresponding decomposition of QIC, consistent with this idea. These definitions are “superposed” variants of the definitions in the previous sections.

Definition 20 *For a protocol Π and a product input distribution $\mu = \mu_X \otimes \mu_Y$, the superposed-classical input information cost of the messages from Alice to Bob (resp. from Bob to Alice) is defined as*

$$\begin{aligned} \text{SCIC}_{A \rightarrow B}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; X | R_Y Y B_i) \\ \left(\text{resp. } \text{SCIC}_{B \rightarrow A}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ even}} I(C_i; Y | R_X X A_i) \right), \end{aligned}$$

the superposed-classical input reverse information cost of the messages from Bob back to Alice (resp. from Alice back to Bob) is defined as

$$\begin{aligned} \text{SCRIC}_{A \leftarrow B}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ even}} I(C_i; X | R_Y Y B_i) \\ \left(\text{resp. } \text{SCRIC}_{B \leftarrow A}(\Pi, \mu) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; Y | R_X X A_i) \right), \end{aligned}$$

the superposed-Holevo information cost from Alice to Bob (resp. from Bob to Alice) is defined as

$$\begin{aligned} \text{SHIC}_{A \rightarrow B}(\Pi, \mu) &= I(X; R_Y Y B_{\text{out}} B') \\ \left(\text{resp. } \text{SHIC}_{B \rightarrow A}(\Pi, \mu) &= I(Y; R_X X A_{\text{out}} A') \right). \end{aligned}$$

Note that S-HIC is indeed the notion of information leakage considered by Ref. [JRS09] in their privacy trade-off for the index function on a uniform distribution.

We now link SCIC and SCRIC to QIC using the following remark. For odd i ,

$$I(C_i : R_X R_Y | Y B_i) = I(C_i : R_Y | Y B_i) + I(C_i : R_X | R_Y Y B_i) \quad (5.6)$$

(we could do similarly for even i with the conditioning instead on $X A_i$). The second term on the right hand side is the superposed-classical input information cost term $I(C_i : R_X | R_Y Y B_i) = I(C_i : X | R_Y Y B_i)$ in round i . For product distributions, it somehow quantifies the amount of information that message C_i contains about X for someone who runs the protocol with the distribution corresponding to Y in a superposition, and also possesses B_i as quantum side-information. The first term does not immediately have such an intuitive interpretation. However, we can rewrite it as $I(C_i : R_Y | Y B_i) = I(C_i : R_Y | R_X X A_i) = I(C_i : Y | R_X X A_i)$ since registers $R_X X A_i$ contains a purification of $\rho_i^{B_i C_i R_Y Y}$. It is then seen to be the superposed-classical input reverse information cost in round i , and hence corresponds to how much information Alice is forgetting about Y if she runs the protocol with the distribution corresponding to X in a superposition, and also possesses A_i as quantum side-information. It follows that $\text{QIC} = \text{SCIC} + \text{SCRIC}$ (Note that this equality also formally holds for non-product distributions if we extend the definitions by using the corresponding CQMI terms).

The Information Flow Lemma can then be used to establish the link with SHIC, noting that for product distributions $I(X; Y R_Y)_{\rho_0} = I(X; Y)_{\rho_i} = 0$, we obtain

$$\text{SHIC}_{A \rightarrow B}(\Pi, \mu) = \text{SCIC}_{A \rightarrow B}(\Pi, \mu) - \text{SCRIC}_{A \leftarrow B}(\Pi, \mu), \quad (5.7)$$

$$\text{SHIC}_{B \rightarrow A}(\Pi, \mu) = \text{SCIC}_{B \rightarrow A}(\Pi, \mu) - \text{SCRIC}_{B \leftarrow A}(\Pi, \mu). \quad (5.8)$$

5.5.2 General Distributions

When considering non-product distributions, if Bob is to run his input in superposition, he should know (at least part of) Alice's input in order to "break the correlations" between their inputs, and allow him to generate the correct superposition consistent with Alice's input. We consider how to do this for running only part of the input in superposition. Notice that this encapsulates and extend both CIC, CRIC, HIC and their superposed variant at once.

Consider tensor product decomposition $X = X_1 \otimes X_2$ of Alice's input and $Y = Y_1 \otimes Y_2$ of Bob's input such that $X_1 Y_1$ and $X_2 Y_2$ are independent, i.e. this gives a product decomposition $XY = X_1 Y_1 \otimes X_2 Y_2$. We can think of Bob running Y_2 in a quantum superposition, and so he also holds the purification R_{X_2} of X_2 in order to generate the correct joint superposition consistent with Alice's input, while being given an actual classical input Y_1 . Alice is then also given a classical input in X_1 (and we can think of X_2 either as a classical input whose classical copy or purification is initially held by Bob, or as a superposition over classical inputs jointly held by Alice and Bob). The corresponding hybrid information costs are defined as follows, with similar definitions for Alice.

Definition 21 For a protocol Π and an arbitrary decomposition $X = X_1 \otimes X_2$, $Y = Y_1 \otimes Y_2$ of the input space, and arbitrary distributions μ_1 on $X_1 Y_1$ and μ_2 on $X_2 Y_2$, when running Π on input distribution $\mu_1 \otimes \mu_2$, the hybrid-classical input information cost of the messages from Alice to Bob (resp. from Bob

to Alice) is defined as

$$\begin{aligned}
\text{HCIC}_{A \rightarrow B}(\Pi, \mu_1, \mu_2) &= \sum_{i \geq 1, i \text{ odd}} I(C_i; X_1 | R_{X_2} R_{Y_2} Y_1 Y_2 B_i) \\
&= \sum_{i \geq 1, i \text{ odd}} I(C_i; X_1 | X_2 R_{Y_2} Y_1 Y_2 B_i) \\
\left(\text{resp. } \text{HCIC}_{B \rightarrow A}(\Pi, \mu_1, \mu_2) \right. &= \sum_{i \geq 1, i \text{ even}} I(C_i; Y_1 | R_{Y_2} R_{X_2} X_1 X_2 A_i) \\
&= \left. \sum_{i \geq 1, i \text{ even}} I(C_i; Y_1 | Y_2 R_{X_2} X_1 X_2 A_i) \right),
\end{aligned}$$

the hybrid-classical input reverse information cost of the messages from Bob back to Alice (resp. from Alice to Bob) is defined as

$$\begin{aligned}
\text{HCRIC}_{A \leftarrow B}(\Pi, \mu_1, \mu_2) &= \sum_{i \geq 1, i \text{ even}} I(C_i; X_1 | R_{X_2} R_{Y_2} Y_1 Y_2 B_i) \\
&= \sum_{i \geq 1, i \text{ even}} I(C_i; X_1 | X_2 R_{Y_2} Y_1 Y_2 B_i) \\
\left(\text{resp. } \text{HCRIC}_{B \leftarrow A}(\Pi, \mu_1, \mu_2) \right. &= \sum_{i \geq 1, i \text{ odd}} I(C_i; Y_1 | R_{Y_2} R_{X_2} X_1 X_2 A_i) \\
&= \left. \sum_{i \geq 1, i \text{ odd}} I(C_i; Y_1 | Y_2 R_{X_2} X_1 X_2 A_i) \right),
\end{aligned}$$

the hybrid-Holevo information cost from Alice to Bob (resp. from Bob to Alice) is defined as

$$\begin{aligned}
\text{HHIC}_{A \rightarrow B}(\Pi, \mu_1, \mu_2) &= I(X'_1; R_{Y_2} Y'_2 B_{\text{out}} B' | Y'_1 X'_2) \\
\left(\text{resp. } \text{HHIC}_{B \rightarrow A}(\Pi, \mu_1, \mu_2) \right. &= I(Y'_1; R_{X_2} X'_2 A_{\text{out}} A' | X'_1 Y'_2) \left. \right).
\end{aligned}$$

Note that by the Information Flow Lemma and the fact that X_1 and X_2 (resp., Y_1 and Y_2) are independent, we get that

$$\text{HHIC}_{A \rightarrow B}(\Pi, \mu_1, \mu_2) = \text{HCIC}_{A \rightarrow B}(\Pi, \mu_1, \mu_2) - \text{HCRIC}_{A \leftarrow B}(\Pi, \mu_1, \mu_2), \quad (5.9)$$

$$\text{HHIC}_{B \rightarrow A}(\Pi, \mu_1, \mu_2) = \text{HCIC}_{B \rightarrow A}(\Pi, \mu_1, \mu_2) - \text{HCRIC}_{B \leftarrow A}(\Pi, \mu_1, \mu_2). \quad (5.10)$$

We then say that Alice does not forget information if the HCRIC from Bob to Alice is 0 for any decomposition of the inputs. More formally, we introduce the following definition.

Definition 22 Given a protocol Π , we say that Alice (resp. Bob) does not forget information in Π if for any decomposition $X = X_1 \otimes X_2$, $Y = Y_1 \otimes Y_2$ of the input space, and any distributions μ_1 on $X_1 Y_1$ and μ_2 on $X_2 Y_2$, it holds that

$$\begin{aligned}
\text{HCRIC}_{A \leftarrow B}(\Pi, \mu_1, \mu_2) &= 0 \\
\left(\text{resp. } \text{HCRIC}_{B \leftarrow A}(\Pi, \mu_1, \mu_2) \right. &= 0 \left. \right).
\end{aligned}$$

We say that protocol Π does not forget information if both Alice and Bob do not forget information in Π .

Remark 23 In particular, if a protocol Π does not forget information, for any input distribution μ , $\text{CRIC}(\Pi, \mu) = 0$, and $\text{QIC}(\Pi, \mu) = \text{HIC}(\Pi, \mu) = \text{CIC}(\Pi, \mu)$.

6 Forgetting Information in Classical Protocols

We considered quantum protocols forgetting classical messages by viewing such messages as part of a quantum register, on which we could apply a reversible quantum operation in order to generate the subsequent message. In the same way, we can consider a reversible classical computation paradigm where classical protocols can forget information. We will show that such an ability does not provide any advantage over protocols in the standard classical information complexity paradigm: for any protocol that can forget information, there exists a protocol that does not forget information with the same input-output behavior, the same amount of communication, and information cost at most that of the protocol that can forget information. In this section, all the protocols we consider are classical.

6.1 Extending the Classical Setting : a New Characterization of IC

Let us begin by deriving some alternative characterization of classical information complexity that will enable easier comparison to the quantum setting. Let us first state some definitions. In the sequel, unless otherwise specified, we denote S_A , S_B , and R_{AB} the random variables corresponding respectively to the private coins of Alice, of Bob, and the public randomness.

Definition 24 A (standard) r -round classical protocol π is defined by the sequence of its message functions such that : for all odd $1 \leq i \leq r$, m_i is a function of $(x, s_A, r_{AB}, m_{<i})$, and for all even $2 \leq i \leq r$, m_i is a function of $(y, s_B, r_{AB}, m_{<i})$.

The randomness of a protocol is contained on the one hand in the inputs (X, Y) and on the other hand in the random coins (S_A, S_B, R_{AB}) .

Definition 25 The (standard) information cost of a protocol π with transcript $\Pi = M_1 \cdots M_r$ on input distribution μ is :

$$IC(\Pi, \mu) = IC_{A \rightarrow B}(\Pi, \mu) + IC_{B \rightarrow A}(\Pi, \mu),$$

where $IC_{A \rightarrow B}(\Pi, \mu) = I(X; \Pi | R_{AB} Y)$, and $IC_{B \rightarrow A}(\Pi, \mu) = I(Y; \Pi | R_{AB} X)$ are respectively the information costs from Alice to Bob and from Bob to Alice, and Π is the sequence of messages.

We generalize the above definitions to the case where there is an additional random variable correlated with the input.

Definition 26 Given a random variable U with distribution μ , we say that a joint random variable UV is an extension of U , or that V extends U , if the marginal of UV on U has distribution μ .

Moreover, we say that V is a copy of U if $\mathbb{P}(U = V) = 1$.

Lemma 27 For any protocol Π , any input distribution μ on XY and any extension $XYX'Y'D$ of XY , where $X'Y'$ are copies of XY , it holds that:

$$\begin{aligned} IC(\Pi, \mu) &= \sum_{i: \text{odd}} I(X'Y'D; M_i | R_{AB} S_B Y M_{<i}) + \sum_{i: \text{even}} I(X'Y'D; M_i | R_{AB} S_A X M_{<i}) \\ &= \sum_i \left(I(X'Y'D; M_i | R_{AB} S_B Y M_{<i}) + I(X'Y'D; M_i | R_{AB} S_A X M_{<i}) \right) \\ &= I(X'Y'D; \Pi | R_{AB} S_B Y) + I(X'Y'D; \Pi | R_{AB} S_A X) \end{aligned} \quad (6.1)$$

Proof. For the first equality, let us consider the right-hand side. In any odd round i , we have :

$$\begin{aligned} &I(X'Y'D; M_i | R_{AB} S_B Y M_{<i}) \\ &= I(Y'; M_i | R_{AB} S_B Y M_{<i}) + I(X'; M_i | Y' R_{AB} S_B Y M_{<i}) + I(D; M_i | X'Y' R_{AB} S_B Y M_{<i}) \\ &= I(X; M_i | R_{AB} S_B Y M_{<i}) \\ &= I(X; M_i | R_{AB} Y M_{<i}), \end{aligned}$$

where we used the following facts. Firstly, $I(Y'; M_i | R_{AB} S_B Y M_{<i}) = 0$, since all the quantities are classical and Y appears in the conditioning. Secondly, $I(D; M_i | X R_{AB} S_B Y M_{<i}) = 0$; indeed, by the Markov property of Π , conditioned on $XY R_{AB} S_B M_{<i}$, M_i is independent of D . Finally, conditioned on either of $XY R_{AB} M_{<i}$ or $Y R_{AB} M_{<i}$, the message M_i generated by Alice is independent of S_B . Similarly, in any even round i , we have :

$$I(X'Y'D; M_i | X R_{AB} S_A M_{<i}) = I(Y; M_i | X R_{AB} M_{<i}).$$

Summing over rounds and using the chain rule of conditional mutual information and Definition 25 yields the first equality.

For the second equality, note that for any odd i

$$\begin{aligned} & I(X'Y'D; M_i | X R_{AB} S_A M_{<i}) \\ &= I(X'; M_i | X R_{AB} S_A M_{<i}) + I(Y'; M_i | X' X R_{AB} S_A M_{<i}) + I(D; M_i | Y' X' X R_{AB} S_A M_{<i}) \\ &= I(Y; M_i | X R_{AB} S_A M_{<i}) \\ &= 0, \end{aligned}$$

in which the last equality follows since M_i is a deterministic function of $X R_{AB} S_A M_{<i}$. Similarly, in any even round i , we have :

$$I(X'Y'D; M_i | R_{AB} S_B Y M_{<i}) = 0.$$

The last equality holds by the chain rule for conditional mutual information. ■

The form (6.1) has a natural interpretation, which we will adopt to define information cost in the reversible classical computation paradigm that we study in the next subsection: it quantifies how much information message M_i in round i contains about any extension of the input, conditional on the information already known at the receiver's side for one term, and on the sender's side for the other term. Since communication in protocols in the reversible classical computation paradigm should be symmetric under time reversal, this will be the natural extension of IC that we will study in that paradigm.

6.2 Reversible Classical Protocols

For notational simplicity, given two registers I and O , we will denote $\mathbf{C}^{I \rightarrow O}$ a reversible circuit taking I as input and outputting in O .

Definition 28 A reversible r -round classical protocol taking X, Y as inputs, with private randomness S_A, S_B and public randomness R_{AB}^A, R_{AB}^B (each player has a copy of the public randomness), and outputting in $A_{out} B_{out}$, is defined by a sequence of reversible circuits : $\mathbf{C}_1^{X S_A R_{AB}^A \rightarrow A_1 M_1}, \mathbf{C}_2^{Y S_B R_{AB}^B M_1 \rightarrow B_2 M_2}, \mathbf{C}_3^{A_1 M_2 \rightarrow A_3 M_3}, \mathbf{C}_4^{B_2 M_3 \rightarrow B_4 M_4}, \dots, \mathbf{C}_r^{A_{r-2} M_{r-1} \rightarrow A' A_{out} M_r}, \mathbf{C}_{r+1}^{B_{r-1} M_r \rightarrow B' B_{out}}$. The circuits are applied in turn by each player to produce the messages M_i (plus some private data A_i or B_i respectively for i odd and i even).

For simplicity, we relabel $A_i = A_{i-1}$ for odd i and $B_i = B_{i-1}$ for even i .

As in the quantum case, we will often require that the players make a copy of their inputs at the outset of the protocol, and run the protocol on these copies while leaving their original input registers unmodified.

Definition 29 A reversible classical protocol is said to be safe if it leaves the input registers unmodified. The safe version of an arbitrary reversible protocol is one in which the players start by making safe copies of their inputs, and then run the protocol on the copies.

We now define a notion of information cost associated to a reversible protocol.

Definition 30 Let Π be a reversible protocol as per Definition 28, let μ be an input distribution on XY , and let $XYX'Y'D$ be any extension of XY , with $X'Y'$ being copies of XY . The reversible information cost of Π on input distribution μ is defined as :

$$\text{RIC}(\Pi, \mu) = \sum_i I(X'Y'D; M_i|B_i) + \sum_i I(X'Y'D; M_i|A_i). \quad (6.2)$$

Note that the sum is over all rounds for both terms of the right-hand side. We first make sure that the above is well-defined, and does not depend on the choice of extension in D . Also, as in the quantum setting, we show that making safe copies does not increase RIC.

Lemma 31 For any reversible protocol Π and input distribution μ , measuring information about $X'Y'$ is sufficient: for any extension $XYX'Y'D$ as above, it holds that

$$\text{RIC}(\Pi, \mu) = \sum_i I(X'Y'; M_i|B_i) + \sum_i I(X'Y'; M_i|A_i). \quad (6.3)$$

Moreover, denoting Π' the safe version of Π , it holds that

$$\text{RIC}(\Pi', \mu) \leq \text{RIC}(\Pi, \mu). \quad (6.4)$$

Proof. By the Markov property of Π , it holds that, conditional on $X'Y'B_i$ or $X'Y'A_i$, M_i and D are independent. The equality (6.3) follows.

For the safe protocol Π' , let $X''Y''$ be safe copies made at the outset to take as input to Π . Then Alice holds X throughout together with A_i , and Bob holds Y throughout together with B_i . It holds that

$$I(X'Y'D; M_i|YB_i) \leq I(XYX'Y'D; M_i|B_i)$$

and

$$I(X'Y'D; M_i|XA_i) \leq I(XYX'Y'D; M_i|A_i).$$

Then, relabeling inputs XY to Π as $X''Y''$, thinking of $XYX'Y'D$ as an extension of input $X''Y''$, and applying (6.3) to Π' and Π leads to (6.4). ■

We thus consider only safe protocols, denote Alice's and Bob's local memory registers as XA_i, YB_i , respectively, and use the following characterization of information cost for these protocols :

$$\text{RIC}(\Pi, \mu) = \sum_i I(X; M_i|YB_{i-1}) + \sum_i I(Y; M_i|XA_{i-1}).$$

For standard IC, we can restrict the sum measuring information about X to odd messages, and the sum measuring information about Y to even messages. The additional terms here are similar to those in the quantum setting and correspond to the cost of forgetting information in a classical reversible paradigm. We want to show that forgetting is in fact useless here. The following lemma will be useful towards this goal.

Lemma 32 The reversible information cost is subadditive : for any two protocols Π_1, Π_2 , and any joint input distribution μ_{12} on $X_1Y_1X_2Y_2$,

$$\text{RIC}(\Pi_1 \otimes \Pi_2, \mu_{12}) \leq \text{RIC}(\Pi_1, \mu_1) + \text{RIC}(\Pi_2, \mu_2),$$

with μ_1 the marginal of μ_{12} on X_1Y_1 , and μ_2 the marginal of μ_{12} on X_2Y_2 .

Proof. Consider an odd round i (Bob is the receiver). The i -th term in RIC of $\Pi_1 \otimes \Pi_2$ on input $X_1X_2Y_1Y_2$ with extension $X_1X_2Y_1Y_2D$ is :

$$\begin{aligned} & I(X_1X_2Y_1Y_2D; M_{1,i}M_{2,i}|B_{1,i-1}B_{2,i-1}) \\ &= I(X_1X_2Y_1Y_2D; M_{1,i}|B_{1,i-1}B_{2,i-1}) + I(X_1X_2Y_1Y_2D; M_{2,i}|M_{1,i}B_{1,i-1}B_{2,i-1}). \end{aligned} \quad (6.5)$$

The i -th term in RIC of Π_1 on inputs X_1Y_1 with extension $X_1Y_1X_2Y_2D$ is :

$$\begin{aligned}
& I(X_1Y_1X_2Y_2D; M_{1,i}|B_{1,i-1}) \\
&= I(X_1Y_1X_2Y_2DA_{2,i-1}M_{2,i-1}B_{2,i-1}; M_{1,i}|B_{1,i-1}) \\
&= I(B_{2,i-1}; M_{1,i}|B_{1,i-1}) + I(X_1Y_1X_2Y_2D; M_{1,i}|B_{1,i-1}B_{2,i-1}) \\
&\quad + I(A_{2,i-1}M_{2,i-1}; M_{1,i-1}|X_1Y_1X_2Y_2DB_{1,i-1}B_{2,i-1}) \\
&\geq I(X_1Y_1X_2Y_2D; M_{1,i}|B_{1,i-1}B_{2,i-1})
\end{aligned}$$

which is the first term in (6.5). Above, the first equality is by first appending uncorrelated registers $S_A S_B R_{AB}^A R_{AB}^B$, and then by invariance of conditional mutual information (CMI) under local reversible processing. The second equality is by the chain rule, and the inequality holds by non-negativity of the CMI.

To obtain the second term in (6.5), let us rewrite the i -th term in RIC of Π_2 on inputs X_2Y_2 with extension $X_1Y_1X_2Y_2D$ as :

$$\begin{aligned}
& I(X_1Y_1X_2Y_2D; M_{2,i}|B_{2,i-1}) \\
&= I(X_1Y_1X_2Y_2DA_{1,i}M_{1,i}B_{1,i}; M_{2,i}|B_{2,i-1}) \\
&= I(M_{1,i}B_{1,i}; M_{2,i}|B_{2,i-1}) + I(X_1Y_1X_2Y_2D; M_{2,i}|M_{1,i}B_{1,i}B_{2,i-1}) \\
&\quad + I(A_{1,i}; M_{2,i}|X_1Y_1X_2Y_2DM_{1,i}B_{1,i}B_{2,i-1}) \\
&\geq I(X_1Y_1X_2Y_2D; M_{2,i}|M_{1,i}B_{1,i-1}B_{2,i-1}),
\end{aligned}$$

with similar arguments as above (and the fact that, since Bob is the receiver, $B_{1,i} = B_{1,i-1}$). We similarly control

$$\begin{aligned}
& I(X_1X_2Y_1Y_2D; M_{1,i}M_{2,i}|A_{1,i}A_{2,i}) \\
&\leq I(X_1X_2Y_1Y_2D; M_{1,i}|A_{1,i}) + I(X_1X_2Y_1Y_2D; M_{2,i}|A_{2,i}). \quad (6.6)
\end{aligned}$$

For any even round i , we obtain similar relationships between the various RIC terms. Summing over rounds yields the conclusion. ■

Theorem 33 *It is possible to simulate any reversible protocol Π by a (standard) protocol Π' that does not forget information without increasing the information or the communication costs.*

Proof. Let Π be a reversible protocol. We assume, without blow up in the information and the communication costs that the protocol makes local copies of the inputs (see Lemma 31). We define Π' as follows: the players run Π , but with each party makes a copy of the message in each round and not further acts on that copy. Then, at round i , we can view the action of the protocol as the combined action of two one-round protocols : Π_1^i , which is a reversible protocol implementing the new message by taking the local registers of the reversible protocol as input, and Π_2^i , which contains the previous messages as side information and does not send any message. Then, we use the subadditivity of RIC (see Lemma 32) on these two protocols. Summing over the rounds, we obtain the desired simulation, since these yield the corresponding RIC of the reversible protocol and its standard version. ■

7 Disjointness: Speed-up for Quantum Protocols needs Forgetting Information

In light of what we saw for classical protocols that forget information, the phenomenon of forgetting information in a quantum protocol might appear useless, or even costly, at first sight. A legitimate question is: *given any safe quantum protocol implementing a classical task, potentially forgetting information,*

is there a protocol that does not forget information and accomplishes the same task at a similar information cost? We give a strong negative answer to this question in the case of the Disjointness problem, showing that the ability to forget information is a necessary quantum feature to obtain any speed-up for computing disjointness.

Recently, the notion of QIC was used by Braverman et al. [BGKK⁺15] to prove an optimal lower bound, up to logarithmic terms, on the bounded-round quantum communication complexity of the disjointness function for n -bit inputs, defined as: for all $x, y \in \{0, 1\}^n$,

$$\text{DISJ}_n(x, y) = \neg (\text{OR}_{i \in [n]}(x_i \text{ AND } y_i)).$$

The authors proved that, for a given number r of rounds of communication, the quantum communication complexity is $\text{QCC}^r(\text{DISJ}_n) \in \tilde{\Omega}(\frac{n}{r} + r)$. We adapt their proof to show that, if we only allow quantum protocols that do not forget information, then the round dependence disappears and we completely lose the quadratic quantum speed-up for computing disjointness. This establishes the fact that, in contrast to the case for classical information cost, the ability to forget information is a necessary feature of quantum protocols.

The high-level idea of the proof in Ref. [BGKK⁺15] can be described as follows. The QIC of any protocol solving DISJ_n is lower bounded by n times the QIC of a protocol solving AND, in which the information is measured with respect to any distribution having zero mass on $(1, 1)$ input. The lower bound on the bounded-round quantum communication for disjointness then follows from the fact that any protocol solving AND must have QIC at least $\tilde{\Omega}(\frac{1}{r})$ on such distributions. This lower bound for AND is in turn proven by reducing back to disjointness, for which they prove that $\text{QIC}(\text{DISJ}_m) \in \Omega(\sqrt{m})$ (for any $m \in \mathbb{N}$), and then constructing a low-information protocol for disjointness by applying coordinate-wise some low-information protocol for AND. The authors were interested in the regime $m \in \tilde{\Theta}(r^2)$. By appropriately subsampling, we can ensure that the QIC of the constructed protocol is close to m times that of the AND protocol on distributions with zero-mass on $(1, 1)$ inputs. The remaining ingredient is a bound on the continuity of QIC in the input distribution.

In fact, this continuity argument is the only place where round complexity comes into play. For the AND function, it states that a r -round protocol Π run on an input distribution with mass w on $(1, 1)$ input has QIC which is $(r \cdot H(w))$ -close to the one of Π run on some input distribution with 0-mass on $(1, 1)$ -input. Note that this factor of r is not present for classical information cost (unless we allow for forgetting information, as in Section 6, in which case it is also there in general) and, at an intuitive level, it can be thought of as arising from the possibility of quantum protocols transmitting r times the same information about the $(1, 1)$ input. In particular, it is not there for quantum protocols that do not forget information, and this is the reason why we can lift the proof of Ref. [BGKK⁺15] to a linear lower bound for such protocols. We formalize this intuition below.

Definition 34 We denote $\mathcal{T}^{r, NF}(f, \varepsilon)$ the set of r -round protocols that solve f with error at most ε and do not forget information as per Definition 22.

Definition 35 We denote $\text{QCC}^{r, NF}(f, \varepsilon)$ (resp. $\text{QIC}^{r, NF}(f, \varepsilon)$) the minimal communication (resp. information) cost achieved by a r -round quantum protocol solving f with error at most ε , and without forgetting information – that is:

$$\text{QCC}^{r, NF}(f, \varepsilon) = \min_{\Pi \in \mathcal{T}^{r, NF}(f, \varepsilon)} \text{QCC}(\Pi), \quad \text{QIC}^{r, NF}(f, \varepsilon) = \inf_{\Pi \in \mathcal{T}^{r, NF}(f, \varepsilon)} \max_{\mu} \text{QIC}(\Pi, \mu).$$

We prove that any protocol solving DISJ_n without forgetting information must have communication $\Omega(n)$.

Theorem 36

$$\text{QCC}^{r, NF}(\text{DISJ}_n, 1/3) \in \Omega(n).$$

First, we can obtain the following result by going over the proof of the corresponding result (Lemma 4.20) in Ref. [BGKK⁺15] and restricting our attention to protocols that do not forget information. The proof, given for completeness, is deferred to the Appendix (see Appendix A.1). We require an additional definition.

Definition 37 We denote $\text{QIC}_0^{r,NF}(\text{AND}, \varepsilon)$ the minimal information cost on input distributions with no support on $(1, 1)$ inputs achieved by a r -round quantum protocol solving AND with error at most ε , and without forgetting information – that is:

$$\text{QIC}_0^{r,NF}(\text{AND}, \varepsilon) = \inf_{\Pi \in \mathcal{T}^{r,NF}(\text{AND}, \varepsilon)} \max_{\mu_0} \text{QIC}(\Pi, \mu_0),$$

in which the maximum is taken over all input distribution satisfying $\mu_0(1, 1) = 0$.

Lemma 38 $\text{QCC}^{r,NF}(\text{DISJ}_n, 1/3) \geq n \cdot \text{QIC}_0^{r,NF}(\text{AND}, 1/3)$.

Furthermore, we adapt the proof of Corollary 4.9 in Ref. [BGKK⁺15] for protocols not forgetting information and obtain the following result. The proof is deferred to the Appendix (see Appendix A.2).

Lemma 39 Suppose we have a protocol Π for AND which does not forget information. Then, for any input distribution μ not concentrated on $(1, 1)$,

$$\text{QIC}(\Pi, \mu) \leq \text{QIC}(\Pi, \mu_0) + H(w)$$

(independently of the number of rounds in Π), where $w = \mu(1, 1) \leq 1/2$, $\mu_0(1, 1) = 0$, $\mu_0(x, y) = \frac{1}{1-w}\mu(x, y)$ for $(x, y) \neq (1, 1)$.

A protocol that does not forget information can be boosted without forgetting information or increasing the number of round, similarly to Lemma 4.15 of Ref. [BGKK⁺15].

Lemma 40 For any function f , any bound on the number of round r and any error parameter $\varepsilon > 0$, the following holds:

$$\text{QIC}^{r,NF}(f, \varepsilon) \leq O(\lg 1/\varepsilon) \text{QIC}^{r,NF}(f, 1/3). \quad (7.1)$$

We make use of the following lower and upper bounds proven in Ref. [BGKK⁺15] (the upper bound follows from the proof of their Lemma 6.1) on the QIC of computing DISJ_m for some parameter $m \in \mathbb{N}$.

Lemma 41 $\text{QIC}(\text{DISJ}_m, 1/3) \in \Omega(\sqrt{m})$.

Lemma 42 For any m , any protocol Π_A computing AND with error $1/m^2$, and any $w \in O(\lg^4(m)/m)$,

$$\text{QIC}(\text{DISJ}_m, 2/m) \leq m \cdot \max_{\mu_w} \text{QIC}(\Pi_A, \mu_w) + o(\sqrt{m}),$$

in which μ_w ranges over all distributions with weight at most w on the $(1, 1)$ -input.

Optimizing over protocols $\Pi_A \in \mathcal{T}^{r,NF}(\text{AND}, 1/m^2)$ in Lemma 42 and combining with Lemma 39, we get, for any $r \geq 1$,

$$\text{QIC}(\text{DISJ}_m, 2/m) \leq m \cdot \left(\text{QIC}_0^{r,NF}(\text{AND}, 1/m^2) + H(w) \right) + o(\sqrt{m}),$$

where the l.h.s. is independent of r . Moreover, by Lemma 41, the left-hand side belongs to $\Omega(\sqrt{m})$, so by further combining with Lemma 40, we can rewrite this as

$$\Omega\left(\frac{1}{\sqrt{m} \lg m}\right) \leq \text{QIC}^{r,NF}(\text{AND}, 1/3). \quad (7.2)$$

The r.h.s. is independent of m , so by fixing m to a large enough constant, we get, for any number of round r ,

$$\text{QIC}^{r,NF}(\text{AND}, 1/3) \in \Omega(1).$$

Hence, by Lemma 38, for any n ,

$$\text{QCC}^{r,NF}(\text{DISJ}_n) \in \Omega(n),$$

which concludes the proof of Theorem 36.

8 Quantum Simulation of Classical Protocols

We now study how to quantumly simulate classical protocols, and how the corresponding QIC behaves. By simulating, we mean that there is a quantum protocol with the same input-output behavior. It turns out that we can always find a quantum simulation with the same information cost as the classical protocol; it is even possible to build this quantum simulation such that it does not forget information.

For the reader's convenience, we deal successively with deterministic protocols, public coin protocols, and protocols with private coins. The latter needs a special care and we give a more detailed explanation on the construction.

Deterministic protocols. Let us consider a classical deterministic (i.e., which does not depend on private or shared randomness) protocol Π . We define the protocol Π_0 which is similar to Π except that Alice and Bob keep local copies of their inputs and of the messages, possibly padding messages with 0's such that the order of speech is known in advance to both and independent of the inputs.

Remark 43 *This might affect the communication cost of the protocol, but does not change the information cost or the input-output behavior.*

Now, we define Π_0^* , the quantum simulation of Π_0 (hence it simulates Π as well). To generate their quantum messages, Alice and Bob run as unitary a classical reversible circuit implementing the protocol in each round, and measure the output registers at the end.

Lemma 44 *The quantum simulation Π_0^* has the same input-output behavior and information cost as the original deterministic protocol Π , and the same communication cost as the padded protocol Π_0 .*

The fact that the information cost is unchanged follows by noticing that each register is classical in HIC, which is equal to the IC of the classical protocol, and also $\text{HIC} = \text{CIC}$ which are then also equal to QIC.

Public Coin Protocols. Let us now consider a classical protocol Π with shared randomness. As above, we define a classical protocol Π_0 similar to Π where the players first make a local copy of the shared randomness, and then pad their messages with 0's such that the order of speech is known in advance to both, independently not only of the input, but also of the randomness.

Then we define the quantum simulation protocol Π_0^* by having Alice and Bob use pure shared entanglement to simulate in a canonical way the shared randomness: make two coherent (quantum), perfectly correlated copies of the random strings, a copy being given to Alice and the other one given to Bob. In this way, if either copy is traced out, the other copy is classical and distributed exactly as the corresponding local copy of the shared randomness.

Viewing a classical protocol with shared randomness as one which is an average over deterministic protocols with fixed random strings, they can then run the corresponding classical deterministic protocol.

Lemma 45 *The input-output behavior and the information cost of the quantum simulation protocol Π_0^* is the same as for the original public coin classical protocol Π , and the communication cost is the same as that of the padded protocol Π_0 .*

Once again, the fact that the information cost is unchanged follows by noticing that each register is classical in HIC, which is equal to the IC of the classical protocol, and also $\text{HIC} = \text{CIC}$ which are then also equal to QIC.

Protocols with Private Randomness. The case of classical protocol that also have private randomness is the most tricky to handle. As a first attempt, the private randomness can be simulated in a way similar to public randomness as described above, except that now both coherent copies of the random strings are given to the same party (the one who owns this private random string in the classical protocol). However, these registers do not look like classical registers in the different information costs, and the above argument for classical protocols with only public randomness cannot be used to argue that the information remains unchanged.

Instead, we use a two-step procedure to obtain a quantum simulation protocol for which we can more easily show that the information cost is maintained. The first step consists in giving a classical simulation protocol of the original protocol in which the private randomness is in some canonical form. In the second step, we simulate quantumly this intermediate classical protocol by applying similar arguments as for classical protocols with only public randomness.

Step 1 : canonical classical simulation. Consider a classical protocol Π . Let us first define a canonical transformation which provides another classical protocol, denoted $\tilde{\Pi}$, in a particular form. For this canonical classical simulation, the idea is to use a lot of fresh private randomness in each round, which directly encodes the distribution over messages in each round in a way which is consistent with the local information (input, shared randomness, and previous messages) of the sender. More precisely, say in round i in Π , Alice is to generate message M_i as a deterministic function of her input X , the shared randomness R_{AB} , her private randomness S_A , and the previous messages $M_{<i} = M_1 \dots M_{i-1}$.

For a given (partial) view $(x, r, m_{<i})$ of Alice at round i (excluding her private randomness), consider the random variable $M_i^{x,r,m_{<i}}$ obtained by "averaging" the private randomness s_A , that is : for any fixed message m ,

$$\Pr[M_i^{x,r,m_{<i}} = m] = \mathbb{P}_{S_A}[m_i(x, S_A, r, m_{<i}) = m].$$

Then the canonical simulation protocol $\tilde{\Pi}$ uses in round i the following random variable (which is given to Alice as fresh private randomness) :

$$T_i^A = \bigotimes_{x,r,m_{<i}} M_i^{A,x,r,m_{<i}},$$

that is, independent copies of the random variable M_i corresponding to each possible local view $(x, r, m_{<i})$. At round i , Alice considers her actual local view $(x, r, m_{<i})$, and sends the message corresponding to $M_i^{A,x,r,m_{<i}}$, that is, the element of her private randomness T_i^A corresponding to her actual local view (the other parts of the private randomness T_i^A are left untouched). Bob acts similarly, with some fresh private randomness T_i^B at each even round i . We denote $T^A = \bigotimes_{i \text{ odd}} T_i^A$ and $T^B = \bigotimes_{i \text{ even}} T_i^B$.

Lemma 46 *In this canonical classical simulation, both the information cost and the communication cost are unchanged : for any input distribution μ ,*

$$\text{IC}(\tilde{\Pi}, \mu) = \text{IC}(\Pi, \mu), \quad \text{CC}(\tilde{\Pi}) = \text{CC}(\Pi).$$

Moreover, the distribution of the joint random variable $XYRM_{\leq n}$ for the whole n -round protocol is also unchanged, and thus so is the input-output behavior.

Step 2 : quantum simulation. We consider a protocol $\tilde{\Pi}_0$ in which the messages of $\tilde{\Pi}$ are padded so that the order of speech is independent of the inputs and both public and private randomness. For the quantum simulation protocol $\tilde{\Pi}_0^*$, private randomness is simulated by giving two coherent local copies to the player and letting him or her work on one of them.

Lemma 47 *The input-output behavior and the information cost of the quantum simulation protocol $\tilde{\Pi}_0^*$ is the same as for the original classical protocol Π with private randomness, and the communication cost is the same as that of the padded protocol $\tilde{\Pi}_0$.*

Proof. We first focus on the CIC term. Consider for instance the third round (Alice is the sender). Dropping the ancilla qubits for brevity, the global quantum state just after Bob receives the third message is then :

$$\rho_3 = \rho^{X, R_X, R_{AB}^A, T^A, M_{\leq 3}^A, M_3^M, Y, R_Y, R_{AB}^B, T^B, M_{\leq 3}^B}$$

where M_i^A and M_i^B denote respectively Alice and Bob's copy of the i -th message, whereas M_3^M is the register that is sent over from Alice to Bob. The third term appearing in CIC is :

$$I(M_3^M : X|Y, R_{AB}^B, T^B, M_{\leq 3}^B), \quad (8.1)$$

where the CQMI is evaluated on the quantum state :

$$\begin{aligned} & \rho^{X, M_3^M, Y, R_{AB}^B, T^B, M_{\leq 3}^B} \\ &= \text{Tr}_{R_X, R_{AB}^A, T^A, M_{\leq 3}^A, R_Y}(\rho_3) \\ &= \sum_{x, y, r, m_{\leq 3}} p(x, y, r, m_{\leq 3}) |x, y, r, m_{\leq 3}\rangle \langle x, y, r, m_{\leq 3}| \otimes \rho^{x, m_3, y, r, T^B, m_{\leq 3}} \end{aligned}$$

for some family of quantum states $(\rho^{x, m_3, y, r, T^B, m_{\leq 3}})_{y, r, m_{\leq 3}}$. For the last equality, we used the fact that the registers $X, M_3^M, Y, R_{AB}^B, M_{\leq 3}^B$ are in a classical state, since the registers $R_X, R_{AB}^A, T^A, M_{\leq 3}^A, R_Y$ are traced out. Furthermore, recall that in the classical protocol $\tilde{\Pi}$, the random variable T^B is defined as :

$$T^B = T_2^B \otimes \left(\bigotimes_{i \geq 2} T_{2i}^B \right) = \left(\bigotimes_{y, r, m_1} M_2^{B, y, r, m_1} \right) \otimes \left(\bigotimes_{i \geq 2} T_{2i}^B \right).$$

In the third round of the quantum protocol, since the registers $R_{AB}^A, M_{\leq 3}^A, Y'$ are already traced out, the quantum state can actually be decomposed as :

$$\begin{aligned} & \rho^{X, M_3^M, Y, R_{AB}^B, T^B, M_{\leq 3}^B} \\ &= \left(\sum_{x, y, r, m_{\leq 3}} p(x, y, r, m_{\leq 3}) |x, y, r, m_{\leq 3}\rangle \langle x, y, r, m_{\leq 3}| \otimes \rho^{x, m_3, y, r, T_2^B, m_{\leq 3}} \right) \otimes \left(\bigotimes_{i \geq 2} \rho^{T_{2i}^B} \right). \end{aligned}$$

Hence, by Lemma 1, the term (8.1) can be written

$$I(M_3^M : X|Y, R_{AB}^B, T^B, M_{\leq 3}^B) = \mathbb{E}_{y, r, m_{\leq 3}} \left[I(M_3^M : X|T_2^B)_{\rho^{X, M_3^M, y, r, T_2^B, m_{\leq 3}}} \right], \quad (8.2)$$

with

$$\rho^{X, M_3^M, y, r, T_2^B, m_{\leq 3}} = \sum_{x, m_3} p(x, m_3) |x, m_3\rangle \langle x, m_3| \otimes \rho^{x, m_3, y, r, T_2^B, m_{\leq 3}},$$

where we use the fact that X and M_3^M are classical since R_X and M_3^A were traced out. The T_2^B is still quantum, but it has a special structure: either M_2^{B, y', r', m'_1} does not correspond to the actual view (y, r, m_1) of Bob, and so it remains in a pure state, or else it corresponds but Alice possesses a coherent copy of M_2^{B, y, r, m_1} , and so Bob's copy is classical once we trace Alice's copy out. It follows that

$$I(M_3^M : X|Y, R_{AB}^B, T^B, M_{\leq 3}^B) = \mathbb{E}_{y, r, m_{\leq 3}} \left[I(M_3^M : X|M_2^{B, y, r, m_1})_{\rho^{X, M_3^M, y, r, M_2^{B, y, r, m_1}, m_{\leq 3}}} \right] \quad (8.3)$$

$$= I(M_3^M : X|M_2^{B, Y, R, M_1}), \quad (8.4)$$

as in $IC(\tilde{\Pi}_0)$.

More generally, consider an odd round i (Bob is the receiver). We can see that, conditioning on the classical part $(y, r, m_{<i})$, all of the quantum registers corresponding to the private randomness T^B on Bob's side fall into two categories :

- either they have never been used (for $j \geq i$, all of T_j^B , or for $j \leq i$, the coordinates of T_j^B which did not correspond to the actual view of Bob at round j), and so remain in a pure state in product form and can be eliminated from the CQMI term,
- or else they have been used but correspond to one of at least some quantum copies of a message previously sent to the other party (the coordinates of T_j^B for $j \leq i$, j odd, corresponding to the local view $(y, r, m_{<j})$ of Bob at round j , hence to a message $M_j^{B,x,r,m_{<j}}$ sent by Bob to Alice). In the CQMI, since one party's registers are traced out, this term of CIC is classical.

Using the chain rule, we see that the i -th term in CIC for the quantum simulation is equal to the i -th term in the information cost of the classical protocol $\tilde{\Pi}$. Similar arguments hold also for any even round. Hence $CIC(\tilde{\Pi}_0^*, \mu) = IC(\tilde{\Pi}, \mu)$. Finally, we can see that $HIC(\tilde{\Pi}_0^*, \mu) = CIC(\tilde{\Pi}_0^*, \mu)$ by using the chain rule in an order so as to be able to apply the above argument to the quantum registers corresponding to private randomness. This implies $CRIC(\tilde{\Pi}_0^*, \mu) = 0$, and $QIC(\tilde{\Pi}_0^*, \mu) = CIC(\tilde{\Pi}_0^*, \mu) = HIC(\tilde{\Pi}_0^*, \mu) = IC(\tilde{\Pi}_0, \mu) = IC(\Pi, \mu)$. ■

Remark 48 In particular for classical protocols, $IC_0(\text{AND}) \in \Omega(1)$ (and then also $CC(\text{DISJ}_n) \in \Omega(n)$) by a standard direct sum argument akin to Lemma 38) follow by using such a quantum simulation that does not forget information and using the result $QIC^{r,NF} \in \Omega(1)$ from the previous section. Surprisingly, the main ingredients going into this proof of the linear lower bound on the classical communication complexity of disjointness are a \sqrt{n} lower bound on the quantum information complexity and a \sqrt{n} upper bound on the quantum communication complexity of disjointness, two \sqrt{n} bounds.

9 Clean Protocols, IP, and Random Functions

9.1 Clean Protocols and Phase Encoding of the Output

The development in this section follows that of Refs [CVDNT99, MW07]. The Information Flow Lemma (see Lemma 3) allows us to translate their arguments about QCC to QIC. The link with IC follows by the general simulation procedure of classical protocols maintaining IC (see Lemma 47).

Given a Boolean function f and any protocol Π computing f with zero-error, we will construct a so-called *clean* protocol Π' also computing f with zero-error, but restoring all registers, except for an output qubit, to their original state. Then, using similar ideas, we define a protocol Π'' where the output is in the phase.

Clean protocol Π' . The action of Π , if we do not trace out the A' , B' registers, is given by the sequence of unitaries $U_1, U_2, \dots, U_r, U_{r+1}$ applied by Alice and Bob in turns. Hence, denoting $U_\Pi = U_{r+1}U_r \dots U_2U_1$, the state at the end of a run of Π on input (x, y) is of the form

$$\begin{aligned} & U_\Pi \left(|x\rangle^X |y\rangle^Y |\psi\rangle^{T_A T_B} \right) \\ &= |x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'}, \end{aligned} \quad (9.1)$$

for some state $|\phi_{xy}\rangle$ depending on both x and y .

We define the protocol Π' as the protocol whose global action is given by $U_\Pi^\dagger CNOT_{B_{out} \rightarrow B'_{out}} U_\Pi$ which uses an additional ancillary qubit $|0\rangle^{B'_{out}}$. In other words, the players start by running Π , which leads to the state (9.1). Then, Bob applies a $CNOT$ gate from B_{out} to B'_{out} , which gives the state

$$\begin{aligned}
& CNOT_{B_{out} \rightarrow B'_{out}} \left(|x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} |0\rangle^{B'_{out}} \right) \\
& = |x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} |f(x, y)\rangle^{B'_{out}}.
\end{aligned}$$

To clean the working registers, the players run the protocol whose action is U_{Π}^{\dagger} , and they obtain

$$\begin{aligned}
& U_{\Pi}^{\dagger} \left(|x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} |f(x, y)\rangle^{B'_{out}} \right) \\
& = |x\rangle^X |y\rangle^Y |\psi\rangle^{T_A T_B} |f(x, y)\rangle^{B'_{out}}.
\end{aligned}$$

So the overall action of Π' is

$$U_{\Pi}^{\dagger} CNOT_{B_{out} \rightarrow B'_{out}} U_{\Pi} \left(|x\rangle^X |y\rangle^Y |\psi\rangle^{T_A T_B} |0\rangle^{B'_{out}} \right) \quad (9.2)$$

$$= |x\rangle^X |y\rangle^Y |\psi\rangle^{T_A T_B} |f(x, y)\rangle^{B'_{out}}. \quad (9.3)$$

Remark 49 Notice that if $QCC_{A \rightarrow B}(\Pi) = a$ (the communication from Alice to Bob), $QCC_{B \rightarrow A}(\Pi) = b$ (the communication from Bob to Alice), then $QCC_{A \rightarrow B}(U_{\Pi}^{\dagger}) = b$, $QCC_{B \rightarrow A}(U_{\Pi}^{\dagger}) = a$; hence $QCC_{A \rightarrow B}(\Pi') = a + b = QCC(\Pi)$. We will later argue something similar for information of zero-error protocols.

Protocol Π'' with output in the phase. We define Π'' similarly to Π' , except that the ancilla register B'_{out} is originally in the state $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, instead of $|0\rangle$ as in Π' . As a consequence, instead of recording $f(x, y)$ in the computational basis of B'_{out} , the players “record” it in the phase. So, after running Π , the players apply $CNOT_{B_{out} \rightarrow B'_{out}}$ to obtain

$$\begin{aligned}
& CNOT_{B_{out} \rightarrow B'_{out}} \left(|x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} |-\rangle^{B'_{out}} \right) \\
& = |x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} (-1)^{f(x, y)} |-\rangle^{B'_{out}}.
\end{aligned}$$

Thus, running U_{Π}^{\dagger} and bringing out the phase, we get

$$\begin{aligned}
& (-1)^{f(x, y)} U_{\Pi}^{\dagger} \left(|x\rangle^X |y\rangle^Y |f(x, y)\rangle^{B_{out}} |\phi_{xy}\rangle^{A_{out} A' B'} |-\rangle^{B'_{out}} \right) \\
& = (-1)^{f(x, y)} |x\rangle^X |y\rangle^Y |\psi\rangle^{T_A T_B} |-\rangle^{B'_{out}}.
\end{aligned}$$

9.2 Relating to $QIC(\Pi, \mu)$

These two protocols, Π' and Π'' , have the same communication cost, and in particular:

$$QCC_{A \rightarrow B}(\Pi'') = QCC_{A \rightarrow B}(\Pi') = QCC(\Pi).$$

We now study their information cost and show the following result.

Proposition 50 For any input distribution μ , Boolean function f , and any zero-error protocol Π for f ,

$$QIC(\Pi, \mu) = QIC_{A \rightarrow B}(\Pi', \mu) = QIC_{A \rightarrow B}(\Pi'', \mu).$$

Proof. It is clear for the first half of Π' and Π'' , when running U_Π forward, that the corresponding information costs are $QIC_{A \rightarrow B}(\Pi, \mu)$. We now argue that for the second half, when running U_Π^\dagger , the corresponding information cost is $QIC_{B \rightarrow A}(\Pi, \mu)$.

Consider first the clean protocol Π' , and view register B'_{out} , containing a copy of $f(x, y)$ while running U_Π^\dagger , as an additional part of a purification register R' for protocol Π : $R' = R'_X R'_Y B'_{out}$. This is justified as follows. We can instead think of B'_{out} as being generated, after running U_Π with $|0\rangle$ in B'_{out} (and thus registers $R'_X R'_Y = R_X R_Y$ purify registers $XY A_i B_i C_i$ for that part), by applying U_f , defined such that $U_f(|x\rangle|y\rangle|0\rangle) = |x\rangle|y\rangle|f(x, y)\rangle$, to the registers $R'_X R'_Y B'_{out}$. Since Π , and hence Π' , is a zero error protocol, and f is a function, the resulting states at that point and when further applying U_Π^\dagger will then be the same in this modified purified view of Π as in the clean protocol Π' , and thus the QIC 's are also the same. But then the QIC 's are also identical to the ones in which we run U_Π and then U_Π^\dagger without making a copy in B'_{out} , since the global states are the same up to unitary U_f being applied to purification registers $R_X R_Y$ and the uncorrelated state $|0\rangle$ in B'_{out} .

We can apply a similar reasoning to the protocol Π'' in which $f(x, y)$ is recorded in the phase, since we can similarly think of $(-1)^{f(x, y)}|-\rangle^{B'_{out}}$ as being part of the reference $R'' = R''_X R''_Y B'_{out}$, with B'_{out} remaining in state $|-\rangle$, and the phase information now being generated by applying $U_{f, phase}$, defined such that $U_{f, phase}(|x\rangle|y\rangle) = (-1)^{f(x, y)}|x\rangle|y\rangle$, on registers $R_X R_Y$.

Finally, notice that if we run U_Π and then U_Π^\dagger without acting on the output, then, using the duality relation $I(R_X R_Y; C_i | Y B_i) = I(R_X R_Y; C_i | X A_i)$, we get that $QIC_{A \rightarrow B}$ of Π' and Π'' in the U_Π^\dagger part is $QIC_{B \rightarrow A}(\Pi, \mu)$, so

$$\begin{aligned} QIC_{A \rightarrow B}(\Pi'', \mu) &= QIC_{A \rightarrow B}(\Pi', \mu) \\ &= QIC_{B \rightarrow A}(\Pi'', \mu) \\ &= QIC_{B \rightarrow A}(\Pi', \mu) \\ &= QIC_{A \rightarrow B}(\Pi, \mu) + QIC_{B \rightarrow A}(\Pi, \mu) \\ &= QIC(\Pi, \mu). \end{aligned}$$

■

9.3 Information Lower Bound

To get a tractable lower bound on $QIC_{A \rightarrow B}(\Pi'', \mu) = QIC(\Pi, \mu)$, we focus on total functions and on product distributions $\mu = \mu_X \otimes \mu_Y$ on XY , and we apply the Information Flow Lemma. Taking the purified view, we have in Π''

$$|\rho'_{f, \mu}\rangle^{X R_X Y R_Y T_A T_B B'_{out}} = \left(\sum_{x, y} (-1)^{f(x, y)} \sqrt{\mu_X(x)} \sqrt{\mu_Y(y)} |xxyy\rangle^{X R_X Y R_Y} \right) |\psi\rangle^{T_A T_B} |-\rangle^{B'_{out}}, \quad (9.4)$$

in which we emphasize the dependance of $|\rho'_{f, \mu}\rangle$ on the function f and the product distribution μ .

Proposition 51 *We have the following lower bound:*

$$QIC(\Pi, \mu) \geq I(R_X; Y R_Y)_{\rho'_{f, \mu}}.$$

Proof. Notice that B'_{out} remains in the pure state $|- \rangle$ throughout, independently of x and y , and we can remove that register from all the information terms below. We have successively the following chain:

$$\begin{aligned}
QIC(\Pi, \mu) &= QIC_{A \rightarrow B}(\Pi'', \mu) \\
&\geq \sum_{i \text{ odd}} I(R_X R_Y; C_i | Y B_i)_{\rho'_{i, \mu}} \\
&\geq \sum_{i \text{ odd}} I(R_X; C_i | R_Y Y B_i)_{\rho'_{i, \mu}} \\
&\geq \sum_{i \text{ odd}} I(R_X; C_i | R_Y Y B_i)_{\rho'_{i, \mu}} - \sum_{i \text{ even}} I(R_X; C_i | R_Y Y B_i)_{\rho'_{i, \mu}} \\
&= I(R_X; Y T_B B'_{out} | R_Y)_{\rho'_{f, \mu}} - I(R_X; Y | R_Y)_{\rho'_{0, \mu}} \tag{9.5} \\
&= I(R_X; Y T_B B'_{out} | R_Y)_{\rho'_{f, \mu}} \tag{9.6} \\
&= I(R_X; Y R_Y)_{\rho'_{f, \mu}}, \tag{9.7}
\end{aligned}$$

where equality (9.5) is obtained by application of the Information Flow Lemma under the form of Corollary 4, with $E_1 = R_X$, $E_2 = R_Y$ in Π'' . Equality (9.6) holds since $\mu = \mu_X \otimes \mu_Y$ is a product distribution, so $|\rho'_{0, \mu}\rangle = \left(\sum_x \sqrt{\mu_X(x)} |xx\rangle^{X R_X}\right) \left(\sum_y \sqrt{\mu_Y(y)} |yy\rangle^{Y R_Y}\right)$. As for equality (9.7), notice first that $I(R_X; Y T_B B'_{out} | R_Y)_{\rho'_{f, \mu}} = I(R_X; Y | R_Y)_{\rho'_{f, \mu}}$ (with the same arguments as above). Moreover, $I(R_X; R_Y)_{\rho'_{f, \mu}} = I(X; Y)_{\rho'_{f, \mu}} = 0$, since this is a classical product state on XY , so $I(R_X; Y | R_Y)_{\rho'_{f, \mu}} = I(R_X; Y R_Y)_{\rho'_{f, \mu}}$. ■

Contrasting $\rho'_{f, \mu}$ to $\rho'_{0, \mu}$, if the $Y R_Y$ registers contain information about the X register, it must be “encoded in the phase” $(-1)^{f(x, y)}$ somehow. Another way to think about it, in the spirit of what was done in [CVDNT99, MW07], is as follow: Alice is given a classical random variable X distributed according to μ_X , Bob locally prepares the pure state $\sum_y \sqrt{\mu_Y(y)} |y\rangle^Y |y\rangle^{R_Y}$ corresponding to μ_Y , and Alice and Bob run Π' . Bob ends up with registers $Y R_Y$ (and $T_B B'_{out}$, which were restored to state $|\psi\rangle^{T_A T_B} |- \rangle^{B'_{out}}$, independent of x) of $\rho'_{f, \mu}$, which we now view as the output of a “noisy” classical-quantum communication channel with input register X , in which the different phases allows to (at least partially, depending on f and μ_Y) distinguish the pure states

$$|\rho'_{f, x, \mu_Y}\rangle^{Y R_Y} = \sum_y (-1)^{f(x, y)} \sqrt{\mu_Y(y)} |y\rangle^Y |y\rangle^{R_Y}, \tag{9.8}$$

corresponding to each x . The (channel) Holevo information $\max_{\mu_X} I(X; Y R_Y)_{\rho'_{f, \mu}}$ is a known asymptotically achievable bound for classical communication over such noisy channels, giving an alternate proof sketch of $I(X; Y R_Y)_{\rho'_{f, \mu}} \leq 2QCC_{A \rightarrow B}(\Pi')$ (also using the optimality of super-dense coding; the factor of two disappear if the messages are classical, and also if we do not allow for pre-shared entanglement in Π).

Now,

$$I(X; Y R_Y)_{\rho'_{f, \mu}} = H(Y R_Y)_{\rho'_{f, \mu}} - H(Y R_Y | X)_{\rho'_{f, \mu}},$$

and

$$H(Y R_Y | X)_{\rho'_{f, \mu}} = \mathbb{E}_X H(Y R_Y)_{\rho'_{f, x, \mu_Y}} = 0,$$

since $|\rho'_{f, x, \mu_Y}\rangle^{Y R_Y}$ is a pure state for each x . Notice that $I(R_X; Y R_Y)_{\rho'_{f, \mu}} = H(Y R_Y)_{\rho'_{f, \mu}}$ only depends on μ_X , μ_Y , and f :

$$\rho_{f, \mu}^{Y R_Y} = \sum_x \mu_X(x) |\rho'_{f, x, \mu_Y}\rangle \langle \rho'_{f, x, \mu_Y}|^{Y R_Y}. \tag{9.9}$$

From Proposition (51) and the discussion above we obtain the following lower bound:

$$QIC(\Pi, \mu) \geq H(YR_Y)_{\rho'_{f,\mu}}. \quad (9.10)$$

9.4 Inner Product function

The case of the Inner Product function was studied using a similar argument in Ref. [CVDNT99]. Let us consider $f(x, y) = IP_n(x, y) = x \cdot y$ on $\lg |X| = \lg |Y| = n$ bits, and take $\mu_X = \mu_Y$ the uniformly random distribution. If Bob is given register R_Y together with Y of ρ'_{f,x,μ_Y} and applies first $(CNOT^{\otimes n})_{Y \rightarrow R_Y}$ and then $H^{\otimes n}$ on Y , he gets, for any fixed x on Alice's side,

$$\begin{aligned} & (H^{\otimes n})^Y (CNOT^{\otimes n})_{Y \rightarrow R_Y} \left(2^{-n/2} \sum_y (-1)^{x \cdot y} |yy\rangle^{YR_Y} \right) \\ &= (H^{\otimes n})^Y \left(2^{-n/2} \sum_y (-1)^{x \cdot y} |y\rangle^Y \right) |0^n\rangle^{R_Y} \\ &= |x\rangle^Y |0^n\rangle^{R_Y}, \end{aligned}$$

since $H^{\otimes n}$ is self-inverse and $H^{\otimes n} |x\rangle = 2^{-n/2} \sum_y (-1)^{x \cdot y} |y\rangle$. By isometric invariance of von Neumann entropy, $H(YR_Y)_{\rho'_{f,\mu}} = H(X') = n$, for X' a classical copy of X . We get

$$QIC(IP_n, \nu, 0) \geq n,$$

with ν the uniform distribution on the inputs. Since we only assumed that Bob can compute the function value in our lower bound, we get a matching upper bound for such protocols, and $QIC(IP_n, \nu, 0) = n$.

9.5 Random Functions

The argument of Ref. [CVDNT99] for the IP function was extended in Ref. [MW07] to the study of arbitrary (total) Boolean function, and in particular to argue about the quantum communication complexity of a random Boolean function. They showed, for ν the uniform distribution on $n + n$ bit inputs (i.e. $\lg |X| = \lg |Y| = n$), that a uniformly random Boolean function f , (a function chosen by picking $f(x, y)$ uniformly at random in $\{0, 1\}$ for each pair (x, y)), satisfies with high probability $H(YR_Y)_{\rho'_{f,\mu}} \geq n(1 - o(1))$, and thus $QCC(f, \nu, 0) \geq n(1 - o(n))$. Moreover, for small enough constant $\varepsilon > 0$, they also show using a continuity argument that $QCC(f, \nu, \varepsilon) \in \Omega(n)$. Thus, most Boolean functions have essentially a linear quantum communication complexity.

We focus on the case $\varepsilon = 0$, and extend their results for QIC of a random function. We use the following result proved in Ref. [MW07]. Here, H_2 is the Rényi entropy of order 2, ν is the uniform distribution on $2n$ -bit strings, and the probability is taken over the random choice of f , also picked uniformly at random in $\{0, 1\}$ for each of the 2^{2n} pairs (x, y) .

Theorem 52

$$\Pr_f [H_2(YR_Y)_{\rho'_{f,\nu}} < (1 - \delta)n] \leq \exp(-(2^{\delta n} - 1)^2/2),$$

where the probability is uniform over Boolean functions of $n + n$ bits.

Since $H_2 \leq H$, we get the following theorem by taking $\delta = 1/\sqrt{n}$ above and using (9.10).

Theorem 53

$$\Pr_f [QIC(f, \nu, 0) < (1 - 1/\sqrt{n})n] \leq \exp(-(2^{\sqrt{n}} - 1)^2/2),$$

where the probability is uniform over Boolean functions of $n + n$ bits.

Hence, except with negligible probability over the choice of a random function f ,

$$QIC(f, \nu, 0) \geq n(1 - o(1)).$$

9.6 Non-Zero Error and Classical Protocols

Using the quantum simulation of classical protocols maintaining the classical IC that we gave in Section 8, the above result also implies a bound for any classical protocol. Moreover, it is known (see Ref. [BGPW13a]) that classical IC is continuous at $\varepsilon = 0$, so we get the following corollary.

Corollary 54

$$\Pr_f \left[\lim_{\varepsilon \rightarrow 0} IC(f, \nu, \varepsilon) < (1 - 1/\sqrt{n})n \right] \leq \exp(-(2^{\sqrt{n}} - 1)^2/2),$$

where the probability is uniform over Boolean functions of $n + n$ bits.

Hence, except with negligible probability over the choice of a random function f , we have

$$\lim_{\varepsilon \rightarrow 0} IC(f, \nu, \varepsilon) \geq n(1 - o(1)).$$

To the best of our knowledge, this is the first proof that $\lim_{\varepsilon \rightarrow 0} IC$ for a random function is essentially n , and not only $\Omega(n)$, which was known at least since the work of Braverman and Weinstein [BW12] proving that discrepancy lower bounds IC (through a compression argument).

It is an important open question to determine whether it also holds that QIC is continuous at $\varepsilon = 0$, which would then imply a similar result in the quantum setting.

Acknowledgments. The authors are very grateful to Anurag Anshu, André Chailloux, Ankit Garg, Iordanis Kerenidis, Ashwin Nayak, and Penghui Yao for many useful discussions. M.L. has been supported by ERC grant QCC. D.T. is supported in part by NSERC, CIFAR, Industry Canada and ARL CDQI program. IQC and PI are supported in part by the Government of Canada and the Province of Ontario. Part of this research was conducted while M.L. was a PhD student with the Institut de Recherche en Informatique Fondamentale, Université Paris Diderot, and while D.T. was a PhD student with the Département d’informatique et de recherche opérationnelle, Université de Montréal and was supported in part by a FRQNT B2 Doctoral research scholarship, and by CryptoWorks21.

References

- [AA05] Scott Aaronson and Andris Ambainis. Quantum search of spatial regions. *Theory Comput.*, 1:47–79, 2005.
- [BGKK⁺15] Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao, and Dave Touchette. Near-optimal bounds on bounded-round quantum communication complexity of disjointness. In *Proc. FOCS’15*, 2015.
- [BGPW13a] Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. From information to exact communication (extended abstract). In *STOC’13—Proceedings of the 2013 ACM Symposium on Theory of Computing*, pages 151–160. ACM, New York, 2013.
- [BGPW13b] Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. Information lower bounds via self-reducibility. In *International Computer Science Symposium in Russia*, pages 183–194. Springer, 2013.
- [BR11] Mark Braverman and Anup Rao. Information equals amortized communication. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science—FOCS 2011*, pages 748–757. IEEE Computer Soc., Los Alamitos, CA, 2011.

- [BW12] Mark Braverman and Omri Weinstein. A discrepancy lower bound for information complexity. In *Approximation, randomization, and combinatorial optimization*, volume 7408 of *Lecture Notes in Comput. Sci.*, pages 459–470. Springer, Heidelberg, 2012.
- [CB97] Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201, 1997.
- [CVDNT99] Richard Cleve, Wim Van Dam, Michael Nielsen, and Alain Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Quantum Computing and Quantum Communications*, pages 61–74. Springer, 1999.
- [dW02] Ronald de Wolf. Quantum communication and complexity. *Theoret. Comput. Sci.*, 287(1):337–353, 2002. Natural computing.
- [DY08] Igor Devetak and Jon Yard. Exact cost of redistributing multipartite quantum states. *Physical Review Letters*, 100(23):230501, 2008.
- [JN14] Rahul Jain and Ashwin Nayak. The space complexity of recognizing well-parenthesized expressions in the streaming model: the index function revisited. *IEEE Transactions on Information Theory*, 66(10):1–23, 2014.
- [JRS03] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A lower bound for the bounded round quantum communication complexity of set disjointness. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 220–229. IEEE, 2003.
- [JRS09] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A property of quantum relative entropy with an application to privacy in quantum communication. *Journal of the ACM (JACM)*, 56(6):33, 2009.
- [Kla02] Hartmut Klauck. On quantum and approximate privacy. In *STACS 2002*, volume 2285 of *Lecture Notes in Comput. Sci.*, pages 335–346. Springer, Berlin, 2002.
- [KLL⁺15] Iordanis Kerenidis, Sophie Laplante, Virginie Lerys, Jérémie Roland, and David Xiao. Lower bounds on information complexity via zero-communication protocols and applications. *SIAM Journal on Computing*, 44(5):1550–1572, 2015.
- [KLLGR15] Iordanis Kerenidis, Mathieu Laurière, François Le Gall, and Mathys Rennela. Privacy in quantum communication complexity. In *Proc. Asian Quantum Information Science Conference*, 2015.
- [KLLGR16] Iordanis Kerenidis, Mathieu Laurière, François Le Gall, and Mathys Rennela. Information cost of quantum communication protocols. *Quantum Inf. Comput.*, 16(3-4):181–196, 2016.
- [KN97] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.
- [KNTSZ07] Hartmut Klauck, Ashwin Nayak, Amnon Ta-Shma, and David Zuckerman. Interaction in quantum communication. *IEEE Transactions on Information Theory*, 53(6):1970–1982, June 2007.
- [MW07] Ashley Montanaro and Andreas Winter. A lower bound on entanglement-assisted quantum communication complexity. In *Proc. ICALP’07*, 2007.
- [NT16] Ashwin Nayak and Dave Touchette. Augmented index and quantum streaming for DYCK(2). *In preparation*, 2016.
- [SSS15] Louis Salvail, Christian Schaffner, and Miroslava Sotáková. Quantifying the leakage of quantum protocols for classical two-party cryptography. *International Journal of Quantum Information*, 13(04):1450041, 2015.
- [Tou15] Dave Touchette. Quantum information complexity. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 317–326. ACM, 2015.

- [Wat15] John Watrous. *Theory of Quantum Information*. 2015. Manuscript of a book, available at <https://cs.uwaterloo.ca/~watrous/>.
- [Wil13] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, New York, 2013.
- [Yao79] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213. ACM, 1979.
- [Yao93] Andrew Chi-Chih Yao. Quantum circuit complexity. In *34th Annual Symposium on Foundations of Computer Science (Palo Alto, CA, 1993)*, pages 352–361. IEEE Comput. Soc. Press, Los Alamitos, CA, 1993.
- [YD09] Jon Yard and Igor Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. *IEEE Transactions on Information Theory*, 55(11):5339–5351, 2009.

A Proofs for Section 7

A.1 Proof of Lemma 38

Let us start by stating two intermediate lemmas that can be proved respectively as Lemma 4.18 and Lemma 4.19 in [BGKK⁺15].

Lemma 55 *For any integers n, r , any $\varepsilon > 0$, and any input distribution μ_0 such that $\mu_0(1, 1) = 0$,*

$$\inf_{\Pi_A \in \mathcal{T}^{r, NF}(\text{AND}, \varepsilon)} \text{QIC}^{r, NF}(\Pi_A, \mu_0) \leq \inf_{\Pi_D \in \mathcal{T}^{r, NF}(\text{DISJ}_n, \varepsilon)} \frac{1}{n} \text{QIC}^{r, NF}(\Pi_D, \mu_0^{\otimes n}).$$

Lemma 56

$$\text{QIC}_0^{r, NF}(\text{AND}, \varepsilon) = \max_{\mu_0 : \mu_0(1, 1) = 0} \inf_{\Pi_A \in \mathcal{T}^{r, NF}(\text{AND}_n, \varepsilon)} \text{QIC}(\Pi, \mu_0).$$

We can now proceed to the proof of Lemma 38.

Proof of Lemma 38. The result is a consequence of the following chain of inequalities:

$$\begin{aligned} \text{QCC}^{r, NF}(\text{DISJ}_n, 1/3) &\geq \text{QIC}^{r, NF}(\text{DISJ}_n, 1/3) \\ &\geq \max_{\mu} \inf_{\Pi_D \in \mathcal{T}^{r, NF}(\text{DISJ}_n, 1/3)} \text{QIC}(\Pi_D, \mu) \\ &\geq \max_{\mu_0 : \mu_0(1, 1) = 0} \inf_{\Pi_D \in \mathcal{T}^{r, NF}(\text{DISJ}_n, 1/3)} \text{QIC}(\Pi_D, \mu_0^{\otimes n}) \\ &\geq \max_{\mu_0 : \mu_0(1, 1) = 0} \inf_{\Pi_A \in \mathcal{T}^{r, NF}(\text{AND}, 1/3)} n \cdot \text{QIC}(\Pi_A, \mu_0) \\ &\geq n \cdot \text{QIC}^{r, NF}(\text{AND}, 1/3). \end{aligned}$$

The first inequality holds since QIC lower bounds QCC, the second since the protocol can now be optimized according to μ , the third since, on the r.h.s. the maximization is over a smaller set of product distributions satisfying $\mu_0(1, 1) = 0$. The fourth is by Lemma 55, and the last is by Lemma 56. ■

A.2 Proof of Lemma 39

As a first step, we show that the second inequality of Lemma 4.7 in Ref. [BGKK⁺15] admits a tighter version for protocols not forgetting information (according to Definition 22).

Lemma 57 (Quasi-convexity in input) *Let $p \in [0, 1]$, and μ_1, μ_2 be two input distribution. Define $\mu = p\mu_1 + (1 - p)\mu_2$. Then the following holds for any protocol Π which does not forget information:*

$$\begin{aligned} \text{QIC}(\Pi, \mu) &\geq p\text{QIC}(\Pi, \mu_1) + (1 - p)\text{QIC}(\Pi, \mu_2), \\ \text{QIC}(\Pi, \mu) &\leq p\text{QIC}(\Pi, \mu_1) + (1 - p)\text{QIC}(\Pi, \mu_2) + H(p), \end{aligned}$$

independently of the number of rounds in Π .

Compared with Lemma 4.7 Ref. [BGKK⁺15], we save a multiplicative factor equals to the number of rounds in front of the term $H(p)$.

Proof of Lemma 57. The first inequality holds by the first inequality of Lemma 4.7 in Ref. [BGKK⁺15]. Let us prove here the second inequality. Since Π does not forget information, by Remark 23, its QIC is equal to its HIC. So it is sufficient to prove the desired inequality with QIC replaced by HIC. Let R be a register holding a purification of ρ_{μ_1} and ρ_{μ_2} . Then, we can purify ρ_μ with two copies S_1, S_2 of a selector reference register, such that

$$|\rho_\mu\rangle^{A_{in}B_{in}RS_1S_2} = \sqrt{p}|\rho_{\mu_1}\rangle^{A_{in}B_{in}R}|1\rangle^{S_1}|1\rangle^{S_2} + \sqrt{1-p}|\rho_{\mu_2}\rangle^{A_{in}B_{in}R}|2\rangle^{S_1}|2\rangle^{S_2}.$$

We can expand the HIC from Alice to Bob as:

$$\begin{aligned} \text{HIC}_{A \rightarrow B}(\Pi, \mu) &= I(X; B_{out}B'|Y)_{\rho_\mu} \\ &= I(XS_1; B_{out}B'|Y)_{\rho_\mu} \\ &= I(S_1; B_{out}B'|Y)_{\rho_\mu} + I(X; B_{out}B'|YS_1)_{\rho_\mu} \\ &\leq H(p) + I(X; B_{out}B'|YS_1)_{\rho_\mu}, \end{aligned}$$

where the first equality is by definition of HIC, the second because $I(S_1; B_{out}B'|XY)_{\rho_\mu} = 0$ by the Markov property of protocols (X, Y and S_1 are all classical here), the third one is by chain rule, and the inequality is by the fact that S_1 is classical and $H(S_1) = H(p)$.

Moreover, since S_1 is a classical register when S_2 is traced out,

$$I(X; B_{out}B'|YS_1)_{\rho_\mu} = pI(X; B_{out}B'|Y)_{\rho_{\mu_1}} + (1 - p)I(X; B_{out}B'|Y)_{\rho_{\mu_2}}.$$

Hence:

$$\text{HIC}_{A \rightarrow B}(\Pi, \mu) \leq H(p) + p\text{HIC}_{A \rightarrow B}(\Pi, \mu_1) + (1 - p)\text{HIC}_{A \rightarrow B}(\Pi, \mu_2).$$

■

Then, we conclude the proof of Lemma 39.

Proof of Lemma 39. Let us denote μ_1 the probability distribution with weight 1 on input $(1, 1)$. Then, we can write:

$$\mu = (1 - w)\mu_0 + w\mu_1, \quad \mu_0 = (1 - w)\mu_0 + w\mu_0.$$

Hence, by Lemma 57:

$$\begin{aligned} \text{QIC}(\Pi, \mu) &\leq (1 - w)\text{QIC}(\Pi, \mu_0) + w\text{QIC}(\Pi, \mu_1) + H(w) \\ &\leq (1 - w)\text{QIC}(\Pi, \mu_0) + H(w) \\ &\leq \text{QIC}(\Pi, \mu_0) + H(w). \end{aligned}$$

■

B The Various Notions of Information Cost

QIC	Definition 2 (see [Tou15])
CIC	Definition 11 (see [KLLGR15])
HIC	Definition 13
CRIC	Definition 14
SCIC, SCRIC, SHIC	Definition 20
HCIC, HCRIC, HHIC	Definition 21
IC	Definition 25
RIC	Definition 30

Table 1: The classical and quantum notions of information cost used in this article.